

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
28 December 2000 (28.12.2000)

PCT

(10) International Publication Number
WO 00/78787 A1

(51) International Patent Classification⁷: **C07H 21/04**,
C07K 1/00, A61K 39/09, 39/38, 38/385, C12N 15/00,
15/09, C12P 21/04, G01N 33/53, 33/569

(74) Agent: **MADSON & METCALF**; 15 West South Temple,
Suite 900, Salt Lake City, UT 84101 (US).

(21) International Application Number: PCT/US00/17082

(22) International Filing Date: 21 June 2000 (21.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/140,084 21 June 1999 (21.06.1999) US

(71) Applicant (for all designated States except US): **UNIVERSITY OF UTAH RESEARCH FOUNDATION** [US/US];
Suite 110, 615 Arapen Drive, Salt Lake City, UT 84108 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ADDERSON, Elisabeth** [US/US]; St. Jude Children's Research Hospital,
Room D 2038, 332 N. Lauderdale, Memphis, TN 38105 (US). **BOHNSACK, John** [US/US]; University of Utah
Health Sciences Center, Dept. of Pediatrics, 50 North Medical Drive, Salt Lake City, UT 84132 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ISOLATED GENES FROM VIRULENT GROUP B <I>STREPTOCOCCUS AGALACTIAE</I>

(57) Abstract: The present invention relates to the identification and prevention of infections by virulent forms of Group B streptococci. Disclosed herein is the identification of two genes, *spb1* and *spb2*, that are specific to virulent type III-3 GBS. Also disclosed herein are diagnostic methods for detecting virulent GBS infections and methods of immunizing a mammal against these bacteria.



WO 00/78787 A1

ISOLATED GENES FROM VIRULENT GROUP B *STREPTOCOCCUS* *AGALACTIAE*

1. FIELD OF THE INVENTION

5 The present invention relates to genes isolated from Group B streptococci (“GBS”). More specifically, the present invention relates to genes that are specific to virulent forms of GBS and methods of using such genes and their products for the diagnosis and treatment of GBS infections.

2. TECHNICAL BACKGROUND

10 Group B streptococci (“GBS”) are a common cause of disease in newborns, pregnant women, and other persons. Common manifestations of these infections include bacteremia, pneumonia, meningitis, endocarditis, and osteoarticular infections. C.J. Baker & M.S. Edwards, *Group B Streptococcal Infections*, in *Infectious Disease of the Fetus and Newborn Infant*, 980-1054 (J.S. Remington & J.O. Klein, eds., 1995); P. Munoz et al., *Arch Int Med* **157**:213-216 (1997).

15 Nearly 3 out of every 1,000 children born are infected with an invasive form of GBS disease. While GBS disease is of great concern in neonates, GBS is also an important pathogen in the general population, in which the incidence of invasive GBS disease is nearly 8 in 100,000. Of these infections, the mortality rate can be as high as 20 30%.

During childbirth, GBS can pass from the mother to the newborn. By one estimate, up to 30% of pregnant women carry GBS at least temporarily in the vagina or rectum without symptoms. Infants born to these women become colonized with GBS during delivery. Baker & Edwards, *supra*. Aspiration of infected amniotic fluid or 25 vaginal secretions allow GBS to gain access to the lungs. Adhesion to, and invasion of, respiratory epithelium and endothelium appear to be critical factors in early onset neonatal infection. Baker & Edwards, *supra*; C.E. Rubens et al., *J Inf Dis* **164**:320-330 (1991). Subsequent steps in infection, such as blood stream invasion and the establishment of metastatic local infections have not been clarified. The pathogenesis of 30 neonatal infection occurring after the first week of life is also not well understood. Gastrointestinal colonization may be more important than a respiratory focus in late onset neonatal disease. Baker & Edwards, *supra*. Considerable evidence suggests that invasion of brain microvascular endothelial cells by GBS is the initial step in the pathogenesis of

meningitis. GBS are able to invade human brain microvascular endothelial cells and type III GBS, which are responsible for the majority of meningitis, accomplish this 2-6 times more efficiently than other serotypes. V. Nizet et al. *Infect Immun* **65**:5074-5081 (1997).

Because GBS is widely distributed among the population and is an important pathogen in newborns, pregnant women are commonly tested for GBS at 26 to 28 weeks of pregnancy. Much of GBS neonatal disease is preventable by administration of prophylactic antibiotics during labor to women who test positive or display known risk factors. However, these antibiotics programs do not prevent all GBS disease. The programs are deficient for a number of reasons. First, the programs can be inefficient. Second, it is difficult to ensure that all healthcare providers and patients comply with the testing and treatment. And finally, if new serotypes or antibiotic resistance emerges, the antibiotic programs may fail altogether. Currently available tests for GBS are inefficient. These tests may provide false negatives. Furthermore, the tests are not specific to virulent strains of GBS. Thus, antibiotic treatment may be given unnecessarily and add to the problem of antibiotic resistance. Although a vaccine would be advantageous, none are yet commercially available.

Traditionally, GBS have been divided into 9 serotypes according to the immunologic reactivity of the polysaccharide capsule. H.M. Blumberg et al., *J Inf Dis* **173**: 365-373 (1996). Serotype III GBS cause 60-70% of all infections and almost all meningitis. Baker & Edwards, *supra*. Type III GBS can be subdivided into three groups of related strains based on the analysis of restriction digest patterns (RDPs) produced by digestion of chromosomal DNA with *Hind* III and *Sse*8387 I. Y. Nagano et al., *J Med Micro* **35**:297-303 (1991); S. Takahashi et al., *J Inf Dis* **177**:1116-1119 (1998). Figure 1 illustrates a comparison of *Hind* III and *Sse*8387 I RDP typing of 62 type III isolates from Salt Lake City, Utah and Tokyo, Japan. Isolates were classified into types based on the similarity of the restriction digest patterns produced by *Hind* III or *Sse*8387 I digestion of chromosomal DNA. The two methods divided the isolates into RDP types containing exactly the same isolates: III-3 contains isolates 1-41, II-2 contains isolates 42-59, and II-1 contains isolates 60-62.

Over 90% of invasive type III GBS neonatal disease in Tokyo, Japan and in Salt Lake City, Utah is caused by bacteria from one of three RDP types, termed RDP type III-3, while RDP type III-2 are significantly more likely to be isolated from vagina than from

blood or CSF. These results suggest that this genetically-related cluster of type III-3 GBS are more virulent than III-2 strains and could be responsible for the majority of invasive type III disease globally.

From the foregoing, it will be appreciated that it would be a significant advancement in the art to provide one or more markers that are specific to virulent type III-3 GBS. It would be a further advancement to provide a method to exploit these markers for clinical identification of virulent type III-3 GBS. It would be a further advancement to provide methods for producing vaccines against type III-3 GBS.

Such compositions and methods are disclosed herein.

3. BRIEF SUMMARY OF THE INVENTION

The present invention relates to markers specific to type III-3 GBS. These markers, the *spb1* and *spb2* gene products (SEQ ID NO: 2 and SEQ ID NO: 4, respectively), are encoded by the *spb1* (SEQ ID NO: 1) and *spb2* (SEQ ID NO: 3) genes. The invention also provides these genes and gene products in substantially purified form.

In certain other embodiments, the present invention relates to recombinant vectors which incorporate the *spb1* gene or other nucleic acid molecules that code for the *spb1* gene product. The recombinant vector may be a plasmid. In certain embodiments, the recombinant vector is a prokaryotic or eukaryotic expression vector. In certain preferred embodiments, the nucleic acid molecule is operably linked to a heterologous promoter and/or other expression control elements, such as heterologous enhancers and polyadenylation sequences.

In certain other embodiments, the present invention relates to recombinant vectors which incorporate the *spb2* gene or other nucleic acid molecules that code for the *spb2* gene product. The recombinant vector may be a plasmid. In certain embodiments, the recombinant vector is a prokaryotic or eukaryotic expression vector. In certain preferred embodiments, the nucleic acid molecule is operably linked to a heterologous promoter and/or other expression control elements.

The present invention also provides host cells comprising the *spb1* and/or *spb2* genes. In other embodiments, a host cell of the present invention comprises nucleic acid molecules that code for the *spb1* and/or *spb2* gene products. The host cell may be a prokaryotic or eukaryotic host cell.

The present invention also relates to diagnostic methods for determining whether a mammal is infected or colonized by virulent GBS. In certain embodiments, a diagnostic method comprises the steps of (1) collecting a bodily fluid or culture from the mammal and (2) analyzing the bodily fluid or culture for the presence or absence of one or more gene products specific to type III-3 GBS, wherein the presence of one or more gene products specific to type III-3 GBS indicates infection or colonization by virulent GBS. The mammal may be a human. Alternatively, the mammal may be a laboratory, domestic, or agricultural animal. The bodily fluid or culture may be any bodily fluid or culture that is typically analyzed for the presence of bacteria. For example, the bodily fluid or culture may be a vaginal or rectovaginal culture. The bodily fluid or culture may also be a throat culture. The bodily fluid or culture may also be an endotracheal tube aspirant, fluid from a bronchoalveolar lavage, or tissue from a lung biopsy. In certain embodiments, the bodily fluid or culture is blood, serum, amniotic fluid, cerebrospinal fluid, or joint fluid. Other sources of material will be apparent to those of skill in the art. In certain embodiments, a diagnostic method of the present invention comprises analyzing a sample for the presence or absence of the *spb1* and/or *spb2* gene product(s).

In certain embodiments of a diagnostic method of the present invention, the polymerase chain reaction ("PCR") is used to identify the presence or absence of the *spb1* and/or *spb2* gene(s). In certain other embodiments, antibodies are used to identify the presence or absence of the *spb1* and/or *spb2* gene products. The antibodies may be monoclonal or polyclonal antibodies.

The present invention also relates to GBS vaccines. In certain embodiments, the present invention provides vaccines comprising the *spb1* gene product, i.e., a protein comprising the amino acid sequence of SEQ ID NO: 2. In certain other embodiments, a vaccine comprises the *spb2* gene product, i.e., a protein comprising the amino acid sequence of SEQ ID NO: 4. In certain preferred embodiments, a vaccine comprises both the *spb1* and *spb2* gene products. The vaccine may include an adjuvant, such as alum. In certain other embodiments, the *spb1* and/or *spb2* gene(s) may be introduced into a mammal using either naked DNA or other gene therapy techniques to induce an immune response against type III GBS.

The present invention further provides methods of immunizing a mammal against GBS infection. In certain embodiments, such methods comprise administering to the

mammal a vaccine comprising an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 2. In certain other embodiments, a method of the present invention comprises administering to the mammal a vaccine comprising an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 4. The vaccine may also contain a mixture of the *spb1* and *spb2* gene products. Vaccines used in the methods of the present invention may further comprise an adjuvant, such as alum.

These and other features and advantages of the present invention will become more fully apparent from the following detailed description.

4. SUMMARY OF THE DRAWINGS

Figure 1 illustrates a comparison of *HindIII* and *Sse83871* RDP typing of 62 type III GBS isolates from Salt Lake City, Utah and Tokyo, Japan. Isolates were classified into types based on the similarity of the restriction digest patterns produced by *HindIII* or *Sse83871* digestion of chromosomal DNA. The two methods divided isolates into RDP types containing exactly the same isolates: III-3 contains isolates 1 - 41, III-2 contains isolates 42 - 59, and III-1 contains isolates 60 - 62.

Figure 2 illustrates a dot blot hybridization of probe 1 with genomic DNA isolated from type III GBS. 10 µg of genomic DNA from each of 62 type III GBS strains was transferred to nylon membrane. Radiolabeled probe 1 hybridized with DNA from all III-3 strains (rows A - D) including the original type III-3 strain (well E1). The probe failed to hybridize with DNA from III-2 strains (F1 - F10, G1 - G7) including the original strain used in the subtraction hybridization (well E10) and III-1 strains (wells H1 - H3). The same pattern of hybridization was observed using clone 3 and 11 probes.

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to the identification and prevention of infections by virulent forms of GBS. The present invention also relates to isolated genes specific to type III-3 GBS. These genes, *spb1* and *spb2*, encode the *spb1* and *spb2* gene products.

The *spb1* and *spb2* genes or other nucleic acid molecules coding for the *spb1* or *spb2* gene products may be incorporated into a recombinant vector using methods known in the art. See, e.g., 1-3 J. Sambrook et al., *Molecular Cloning: A Laboratory Manual*

(2d ed. 1989). Recombinant vectors include any genetic element, such as a plasmid, phage, transposon, cosmid, chromosome, virus, etc., that is capable of replication when associated with the proper control elements and that can transfer gene sequences between cells. Thus, the term includes cloning and expression vehicles.

A nucleic acid molecule of the present invention may be operably linked to expression control sequences, such as heterologous promoters. Examples include, but are not limited to, viral promoters such as the SV40 early promoter and the CMV immediate early promoter region, bacterial promoters, mammalian promoters, inducible promoters, synthetic promoters, hybrid promoters, and the like. Other expression control sequences are known in the art and include polyadenylation signals, transcription termination sequences, upstream regulatory domains, origins of replication, internal ribosome entry sites ("IRES"), and enhancers. These expression control sequences collectively provide for the replication, transcription and translation of a coding sequence in a recipient cell. Not all of these control sequences need always be present in a recombinant vector, so long as the selected coding sequence is capable of being replicated, transcribed and translated in an appropriate host cell.

Recombinant vectors can be constructed to include selectable markers. Suitable markers include genes which confer antibiotic resistance or sensitivity, or impart color, or change the antigenic characteristics when host cells which have been transfected with the recombinant vectors are grown in an appropriate selective medium. Suitable markers are known to those of skill in the art.

The discovery of type III-3 GBS-specific gene products will allow clinicians to diagnose and treat infection and colonization with virulent GBS. For example, hybridization-based assays may be used to determine whether a GBS isolate is type III-3. Figure 2 illustrates the results of hybridization assays with a III-3-specific probe.

Likewise, PCR may be used to detect the presence or absence of either the *spb1* gene or the *spb2* gene (or both) in samples from patients. PCR methods are described generally in C.R. Newton & A. Graham, *PCR* (2nd. ed. 1997); *PCR: Essential Techniques* (J.F. Burke ed., 1996). Patients who are infected with type III-3 GBS may then receive appropriate antibiotic therapy.

Antibodies may also be used to detect the presence or absence of the *spb1* and/or *spb2* gene product(s). Methods for preparing both monoclonal and polyclonal antibodies are described in, e.g., E. Harlow & D. Lane, *Antibodies: A Laboratory Manual* (1988).

The present invention also relates to methods for producing type III-3 GBS vaccines. *See generally Vaccine Protocols* (A. Robinson, G.H. Farrar & C.N. Wiblin eds. 1996). In certain embodiments, the *spb1* and/or *spb2* gene product(s) may be used to immunize against GBS. These gene products may be produced in large quantities using techniques that are known in the art. For example, the appropriate gene or genes may be linked to a prokaryotic promoter and expressed in bacteria. The gene products may then be purified using conventional techniques and used to vaccinate at-risk individuals. Alternatively, the appropriate gene or genes may be linked to a eukaryotic promoter and enhancer (e.g., yeast, baculovirus, SV40, etc.) and expressed in an appropriate cell type. The gene products may then be purified using conventional techniques.

The *spb1* and/or *spb2* gene products, or immunogenic fragments thereof, may stimulate an immune response when administered to a host. Recombinantly produced proteins are especially desirable, as they can be produced in large amounts and purified. Furthermore, recombinantly produced proteins may be engineered to maximize desirable activities and to minimize unwanted effects.

The recombinantly produced *spb1* and/or *spb2* gene products may be used as carrier proteins for a polysaccharide-protein or oligosaccharide-protein conjugate vaccine. *See, e.g.,* R. Schneerson, et al., *Infect Immun* 60:3528-3532 (1992) (describing a *Pneumococcus*-pertussis toxin conjugate vaccine). For example, *Haemophilus influenzae* B vaccines have been produced by conjugating a tetanus toxoid; a *Corynebacterium* toxin, CRM₁₉₇ (which is a mutant diphtheria toxin); and a *Neisseria* outer membrane protein. Oligo- and polysaccharides from GBS might be used in a vaccine. Oligosaccharide- and polysaccharide-protein conjugates alter the immunological properties of the polysaccharide or oligosaccharide and may improve the immune response.

An adjuvant may be used to enhance the immune response to a vaccine containing the *spb1* and/or *spb2* gene products. An adjuvant is any substance that enhances the immune response to an antigen. Without being bound by any particular theory, adjuvants may act by retaining the antigen locally near the site of administration to produce a depot

effect, facilitating the slow, sustained release of the antigen to cells of the immune system. Adjuvants may also attract cells of the immune system. Aluminum hydroxide and aluminum phosphate (collectively and commonly referred to as “alum”) are routinely used as adjuvants in human and veterinary vaccines. Currently, alum is the only adjuvant licensed for human use, although a number of experimental adjuvants are being tested.

5 The *spb1* and/or *spb2* gene(s) may also be introduced into a mammal using either naked DNA or other gene therapy techniques to induce an immune response against virulent GBS.

10 All publications, patents, and patent applications cited in this application are hereby incorporated by reference. U.S. Patent Application Serial No. 60/140,084 is hereby incorporated by reference in its entirety.

6. EXAMPLES

15 The following examples are given to illustrate several embodiments which have been made within the scope of the present invention. It is to be understood that these examples are neither comprehensive nor exhaustive of the many types of embodiments which can be prepared in accordance with the present invention.

Example 1 - Isolation of the *spb1* and *spb2* genes

20 Bacterial factors that contribute to the increased virulence of III-3 strains can be identified by characterizing the differences between the genetic composition of III-3 and III-2 strains. Such genetic differences will be found in the bacterial chromosomes since these strains do not contain plasmids. Takahashi et al., *supra*.

25 To identify genes present in virulent type III-3 GBS strains and not in the avirulent type III-2 strains, a modification of the technique described by Lisitsyn et al., *Science* **259**:946-951 (1993), was used. High molecular weight genomic DNA from an invasive RDP type III-3 GBS strain (strain 874391) and a colonizing (“avirulent”) RDP type III-2 strain (strain 865043) were prepared by cell lysis with mutanolysin and Proteinase K digestion. Y. Nagano et al., *supra*. For genetic subtraction, genomic DNA from both strains was digested with *Taq* I. *Taq* I-digested DNA from the virulent strain was mixed with two complementary oligonucleotides, *Taq*A (5'-CTAGGTGGA-
30 TCCTTCGGCAAT-3' (SEQ ID NO: 5)) and *Taq*B (5'-CGATTGCCGA-3' (SEQ ID NO: 6)), heated to 50°C for 5 minutes, then allowed to cool slowly to 16°C in T4 ligase

buffer. Oligonucleotides were ligated to the virulent strain DNA by incubation with 20 units of T4 ligase at 16°C for 12 hours. After ligation, 500 ng of DNA from the virulent strain, with ligated linkers, and 40 µg of DNA from the avirulent strain, without linkers, was mixed together, denatured by heating, and hybridized at 68°C for 20 hours.

Ten percent of the resulting hybridization mixture was incubated with *Taq* DNA polymerase and dNTPs to fill in the ends of annealed virulent strain DNA. The hybridized DNA was amplified by *Taq* DNA polymerase for 10 cycles using the *TaqA* oligonucleotide as the forward and reverse amplification primer. After amplification, single stranded products remaining after amplification were digested with mung bean nuclease. Twenty percent of the resulting product was then reamplified for 20 cycles. This process of subtraction followed by PCR amplification results in enhanced amplification of DNA segments from the III-3 strains that do not hybridize with DNA segments from the III-2 strains.

A total of four cycles of subtraction and amplification were carried out using successively smaller quantities of III-3 specific PCR products. Two pairs of oligonucleotides were used for subtraction, which were alternated with successive rounds of subtraction-amplification. The oligos were *TaqA* (SEQ ID NO: 5) and *TaqB* (SEQ ID NO: 6) (the first pair) and *TaqE* (5'-AGGCAACTGTGCTAACCGAGGGAAT-3' (SEQ ID NO: 7)) and *TaqF* (5'-CGATTCCCTCG-3' (SEQ ID NO: 8)) (the second pair). The final amplification products were ligated into pBS KS+ vector and transformed into competent XL1-Blue strain *E. coli*. Thirteen clones were randomly selected for analysis. Cross hybridization studies of the 13 inserts revealed that 6 were unique. These probes were used in slot and dot blot experiments to determine whether subtraction was successful and to identify probes hybridizing with all III-3 strains. Each of the 6 unique probes hybridized with the parental III-3 virulent strain, while none of the probes hybridized with the avirulent III-2 strains. Three of the amplified sequence tags (clones 1, 3 and 11) hybridized with genomic DNA from all 62 type III-3 isolates, but did not hybridize with DNA prepared from the III-2 and III-1 isolates.

Figure 2 illustrates a dot blot hybridization of type III GBS genomic DNA hybridized with a clone 1 probe. 10 µg of genomic DNA from each of 62 type III GBS strains was transferred to nylon membrane. Radiolabeled clone 1 probe hybridized with DNA from all III-3 strains (rows A-D) including the original type III-3 strain (well E-1).

The probe failed to hybridize with DNA from III-2 strains (F1-F10, G1-G7) including the original strain used in the subtraction hybridization (well E10) and III-1 strains (wells H1-H3). The same pattern of hybridization was observed using clone 3 and 11 probes. These data demonstrate the feasibility of identifying genes unique to III-3 strains by this method of PCR-based subtractive hybridization, and further support the validity of the RDP

5 typing for identifying genetically-related type III GBS.

The three GBS type III-3-specific sequence tags are short (130-360 bp). To obtain additional sequence information, a genomic GBS III-3 library was constructed. High molecular weight GBS genomic DNA was partially digested with *Bgl* II and cloned into λ FIX II phage vector. Phage were packaged and the library, consisting of 1.7×10^5 recombinant phage containing inserts with a mean size of about 18 kb (totaling approximately 3×10^9 bp), was amplified once. Multiple plaques hybridizing with each of the III-3 GBS-specific probes were purified for further characterization.

10 Three overlapping genomic clones hybridizing with probe 1 were identified, with approximate sizes of 9, 22, and 23 kb. Since the boundaries of GBS III-3 specific DNA are not known, smaller fragments were subcloned and the DNA was verified present in virulent GBS strains before proceeding with further characterization. The first segment examined is a 6.4 kb *Sal* I-*Bgl* II fragment. This genomic DNA is present in all RDP type III-3 strains and in none of 17 RDP type III-2 strains.

20 Over 90% of this genomic DNA fragment has been sequenced and found to contain 5 open reading frames ("ORFs"). Three of these are likely to be authentic genes. They feature ATG start sites, are preceded by potential ribosomal binding sites and, in two cases, stop codons are followed by palindromic sequences that may represent transcriptional terminators. They are relatively short, however, and do not show significant homology at the nucleic acid or amino acid level with sequences registered with Genbank or the available bacterial genome databases.

25 The two other ORFs appear to be more obvious candidates for virulence genes. The *spb1* gene includes a 1509 bp ORF that is preceded by a potential ribosomal binding site 10 bases upstream from an ATG start codon. The predicted protein (502 amino acids and Mr 53,446) has the characteristics of a cell wall-bound protein. The N-terminus of the predicted protein is a hydrophilic, basic stretch of 6 amino acids followed by a 23 amino acid hydrophobic, proline-rich core, consistent with a signal peptide. The

hydrophilic mature protein terminates in a typical LPXTG (SEQ ID NO: 9) domain that immediately precedes a hydrophobic 20 amino acid core and a short, basic hydrophilic terminus.

The nucleotide sequence is not homologous to sequences of other known bacterial genes. The translated amino acid sequence, however, shares segmental homology with a number of characterized proteins, including the fimbrial type 2 protein of *Actinomyces naeslundii* (27% identity over 350 amino acids) and the fimbrial type 1 protein of *Actinomyces viscosus* (25% homology over 420 amino acids), the T6 surface protein of *S. pyogenes* (23% identity over 359 amino acids), and the *hsf* (27% identity over 260 amino acids) and HMW1 adhesins (25% identity over 285 amino acids) of *Haemophilus influenzae*. See M.K. Yeung & J.O. Cisar, *J Bacteriol* **172**:242-2468 (1990); O. Schneewind, et al., *J Bacteriol* **172**:3310-3317 (1990); J.W. St. Geme III, et al., *J Bacteriol* **178**:6281-6287 (1996); J.W. St. Geme III, *Infect Immun* **62**: 3881-3889 (1994). The function of the *S. pyogenes* T6 protein is unknown. Each of the other homologues plays a role in bacterial adhesion and/or invasion.

The second ORF, *spb2*, terminates 37 bp upstream from *spb1* and is in the same transcriptional orientation. This 1692 bp ORF has a deduced amino acid sequence of 563 residues and Mr 64,492. It shares 50.5% nucleic acid identity and 20.7% amino acid identity with *spb1*. Conservation is highest in the carboxy-terminal regions, including a shared LPSTGG (SEQ ID NO: 10) motif. In contrast to *spb1*, *spb2* does not have an obvious signal sequence. Its secretion may be mediated by carboxy-terminal recognition sequences or by accessory peptides. T. Michiels, et al., *Infect Immun* **58**:2840-2849 (1990). The deduced amino acid sequence of *spb2* is also homologous with *S. pyogenes* T6 and *Actinomyces naeslundii* proteins, and to *Listeria monocytogenes* internalin A (22% identity over 308 amino acids)—again, proteins important in adhesion and invasion.

Neither of the predicted gene products has the repetitive structure of previously described GBS surface proteins such as the C and Rib proteins or of *L. monocytogenes* internalin family members. L.C. Madoff et al., *Infect Immun* **59**:2638-2644 (1991); J. Gaillard, et al., *Cell* **65**: 1127-1141 (1991). Hybridization of the originating strain 874391 genomic DNA with an *spb1* probe under low stringency conditions identifies a single band in *EcoR* I, *Sal* I and *Sst* I digests (data not shown), suggesting that a single

copy of *spb1* is present in this strain and that *spb1* is not a member of a significantly homologous “family” of genes.

Example 2 - Adherence and Invasion Assays Using *spb1*⁻ Bacteria

Genomic subtraction identified a 1509 bp ORF, *spb1*, which is present in virulent RDP type III-3 GBS and not in RDP type III-2 strains. The predicted 53 kD protein product has the characteristics of a typical gram positive cell-wall bound protein. The nucleic acid sequence of *spb1* is not homologous to sequences of other known bacterial genes, however the translated amino acid sequence shares segmental homology with several characterized adhesins, including *Actinomyces* fimbrial proteins and *H. influenzae* HMW1, suggesting that Spb1 might contribute to GBS adhesion or invasion. A *spb1*⁻ isogenic deletion mutant GBS strain was created by homologous recombination and the ability of the *spb1*⁻ mutant to adhere to and invade A549 respiratory epithelial cells was determined. Compared to the wild type strain, the number of *spb1*⁻ bacteria adherent to A549 monolayers was reduced by 60.0% (p<0.01) and the number of intracellular invading bacteria was reduced by 53.6% (p<0.01). Without being bound by any particular theory, it appears that Spb1 may contribute to the pathogenesis of GBS pneumonia and bacterial entry into the bloodstream.

CLAIMS:

1. An isolated nucleic acid molecule comprising nucleotides which code for the amino acid sequence of SEQ ID NO: 2.
2. A recombinant vector comprising the nucleic acid molecule of claim 1.
3. The recombinant vector of claim 2, wherein said recombinant vector is a plasmid.
- 5 4. The recombinant vector of claim 2, wherein said recombinant vector is a prokaryotic or eukaryotic expression vector.
5. The recombinant vector of claim 2, wherein the nucleic acid molecule is operably linked to a heterologous promoter.
6. A host cell comprising the nucleic acid molecule of claim 1.
- 10 7. The host cell of claim 6, wherein the host cell is a eukaryotic host cell.
8. The host cell of claim 6, wherein the host cell is a prokaryotic host cell.
9. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1.
10. A recombinant vector comprising the nucleic acid molecule of claim 9.
- 15 11. The recombinant vector of claim 10, wherein said recombinant vector is a plasmid.
12. The recombinant vector of claim 10, wherein said recombinant vector is a prokaryotic or eukaryotic expression vector.
13. The recombinant vector of claim 10, wherein the nucleic acid molecule is operably linked to a heterologous promoter.
- 20 14. A host cell comprising the nucleic acid molecule of claim 9.
15. The host cell of claim 14, wherein the host cell is a eukaryotic host cell.
16. The host cell of claim 14, wherein the host cell is a prokaryotic host cell.
17. An isolated nucleic acid molecule comprising nucleotides which code for the amino acid sequence of SEQ ID NO: 4
- 25 18. A recombinant vector comprising the nucleic acid molecule of claim 17.
19. The recombinant vector of claim 18, wherein said recombinant vector is a plasmid.
20. The recombinant vector of claim 18, wherein said recombinant vector is a prokaryotic or eukaryotic expression vector.
- 30

21. The recombinant vector of claim 18, wherein the nucleic acid molecule is operably linked to a heterologous promoter.
22. A host cell comprising the nucleic acid molecule of claim 17.
23. The host cell of claim 22, wherein the host cell is a eukaryotic host cell.
24. The host cell of claim 22, wherein the host cell is a prokaryotic host cell.
- 5 25. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 3.
26. A recombinant vector comprising the nucleic acid molecule of claim 25.
27. The recombinant vector of claim 26, wherein said recombinant vector is a plasmid.
- 10 28. The recombinant vector of claim 26, wherein said recombinant vector is a prokaryotic or eukaryotic expression vector.
29. The recombinant vector of claim 26, wherein the nucleic acid molecule is operably linked to a heterologous promoter.
30. A host cell comprising the nucleic acid molecule of claim 25.
- 15 31. The host cell of claim 30, wherein the host cell is a eukaryotic host cell.
32. The host cell of claim 30, wherein the host cell is a prokaryotic host cell.
33. A method of immunizing a mammal against Group B streptococci infection, said method comprising administering to the mammal a vaccine comprising an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 2.
- 20 34. The method of claim 33, wherein the vaccine further comprises an adjuvant.
35. The method of claim 34, wherein the adjuvant comprises alum.
36. The method of claim 33, wherein the vaccine further comprises an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 4.
- 25 37. The method of claim 36, wherein the vaccine further comprises an adjuvant.
38. A method of immunizing a mammal against Group B streptococci infection, said method comprising administering to the mammal a vaccine comprising an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 4.
- 30 39. The method of claim 38, wherein the vaccine further comprises an adjuvant.

40. The method of claim 39, wherein the adjuvant comprises alum.

41. A diagnostic method for determining whether a mammal is infected or colonized by virulent Group B streptococci (GBS), said method comprising the steps of:

- (a) collecting a bodily fluid or culture from the mammal;
- (b) analyzing the bodily fluid or culture for the presence or absence of one or more gene products specific to type III-3 GBS;

wherein the presence of one or more gene products specific to type III-3 GBS indicates infection or colonization by virulent GBS.

42. The diagnostic method of claim 41, wherein the one or more gene products specific to type III-3 GBS comprise a protein, said protein comprising the amino acid sequence of SEQ ID NO: 2.

43. The diagnostic method of claim 41, wherein the one or more gene products specific to type III-3 GBS comprise a protein, said protein comprising the amino acid sequence of SEQ ID NO: 4.

44. The diagnostic method of claim 41, wherein the one or more gene products specific to type III-3 GBS comprise a first protein comprising the amino acid sequence of SEQ ID NO: 2 and a second protein comprising the amino acid sequence of SEQ ID NO: 4.

45. The diagnostic method of claim 41, wherein the mammal is a human.

46. The diagnostic method of claim 41, wherein the bodily fluid or culture is a vaginal or rectovaginal culture.

47. The diagnostic method of claim 41, wherein the bodily fluid or culture is a throat culture.

48. The diagnostic method of claim 41, wherein the bodily fluid or culture is blood, serum, amniotic fluid, cerebrospinal fluid, or joint fluid.

49. The diagnostic method of claim 41, wherein the analysis step comprises using polymerase chain reaction (PCR) to identify the presence or absence of one or more gene products specific to type III-3 GBS.

50. The diagnostic method of claim 49, wherein the analysis step comprises using PCR to determine the presence or absence of the *spb1* gene product.

51. The diagnostic method of claim 49, wherein the analysis step comprises using PCR to determine the presence or absence of the *spb2* gene product.

52. The diagnostic method of claim 41, wherein the analysis step comprises using antibodies to identify the presence or absence of one or more gene products specific to type III-3 GBS.

53. The diagnostic method of claim 52, wherein the antibodies are monoclonal antibodies.

5 54. The diagnostic method of claim 53, wherein the monoclonal antibodies are specific for the *spb1* gene product.

55. The diagnostic method of claim 53, wherein the monoclonal antibodies are specific for the *spb2* gene product.

10 56. An isolated and purified protein comprising the amino acid sequence of SEQ ID NO: 2.

57. A vaccine for immunizing a mammalian host against virulent Group B streptococci infection, said vaccine comprising the protein of claim 56.

58. The vaccine of claim 57, further comprising an adjuvant.

59. The vaccine of claim 58, wherein the adjuvant comprises alum.

15 60. The vaccine of claim 57, wherein the protein is conjugated to a bacterial polysaccharide or oligosaccharide.

61. An isolated and purified protein comprising the amino acid sequence of SEQ ID NO: 4.

20 62. A vaccine for immunizing a mammalian host against virulent Group B streptococci infection, said vaccine comprising the protein of claim 61.

63. The vaccine of claim 62, further comprising an adjuvant.

64. The vaccine of claim 63, wherein the adjuvant comprises alum.

65. The vaccine of claim 62, further comprising an isolated and purified protein comprising the amino acid sequence of SEQ ID NO: 2.

25 66. The vaccine of claim 65, further comprising an adjuvant.

67. The vaccine of claim 62, wherein the protein is conjugated to a bacterial polysaccharide or oligosaccharide.

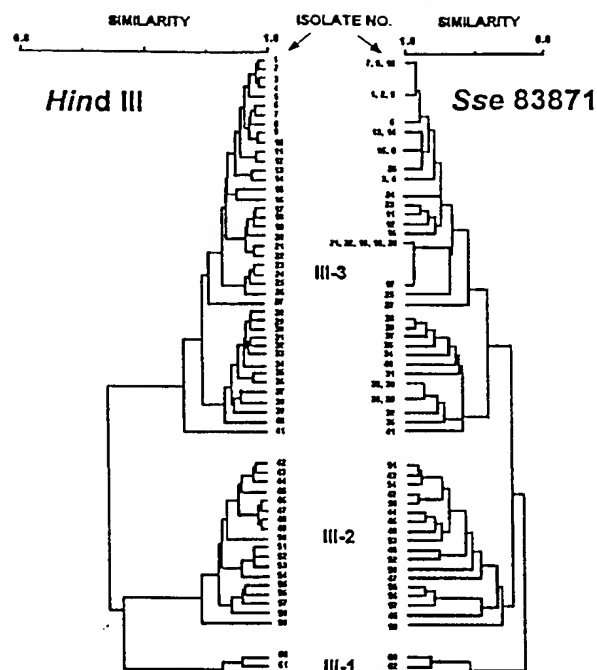


Figure 1. Type III GBS RDP types. Comparison of *Hind*III and *Sse*83871 RDP typing of 62 type III GBS isolates from Salt Lake City and Tokyo. Isolates were classified into types based on the similarity of the restriction digest patterns produced by *Hind*III or *Sse*83871 digestion of chromosomal DNA. The two methods divided the isolates into RDP types containing exactly the same isolates: III-3 contains isolates 1 - 41, III-2 contains isolates 42 - 59, and III-1 contains isolates 60-62.

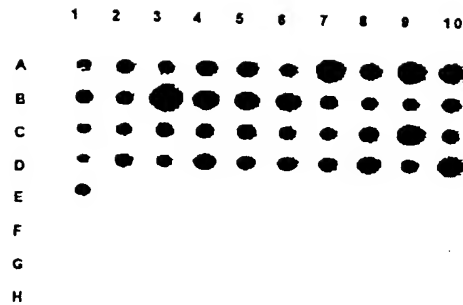


Figure 2. RDP type III-3 specific probes. Dot blot hybridization of probe 1 with genomic DNA isolated from type III GBS. 10 ug of genomic DNA from each of 62 type III GBS strains was transferred to nylon membrane. Radiolabeled probe 1 hybridized with DNA from all III-3 strains (rows A-D) including the original type III-3 strain (well E-1). The probe failed to hybridize with DNA from III-2 strains (F1-F10, G1-7) including the original strain used in the subtraction hybridization (well E 10) and III-1 strains (wells H1-3). The same pattern of hybridization was observed using clone 3 and 11 probes.

SEQUENCE LISTING

<110> Adderson, Elisabeth E.
Bohnsack, John F.
University of Utah Research Foundation

<120> Isolated Genes from Virulent Group B Streptococcus
agalactiae

<130> 1321.2.29

<140>
<141>

<150> 60/140,084
<151> 1999-06-21

<160> 10

<170> PatentIn Ver. 2.1

<210> 1
<211> 1509
<212> DNA
<213> Streptococcus agalactiae

<220>
<221> CDS
<222> (1)..(1509)

<400> 1
atg aaa aag aaa atg att caa tcg ctg tta gtg gcg agt tta gca ttt 48
Met Lys Lys Lys Met Ile Gln Ser Leu Leu Val Ala Ser Leu Ala Phe
1 5 10 15
ggg atg gct gta tca cca gtt acg ccg ata gct ttt gcc gct gag aca 96
Gly Met Ala Val Ser Pro Val Thr Pro Ile Ala Phe Ala Ala Glu Thr
20 25 30
ggg aca att aca gtt caa gat act caa aaa ggc gca acc tat aaa gca 144
Gly Thr Ile Thr Val Gln Asp Thr Gln Lys Gly Ala Thr Tyr Lys Ala
35 40 45
tat aaa gtt ttt gat gca gaa ata gat aat gca aat gta tct gat tcg 192
Tyr Lys Val Phe Asp Ala Glu Ile Asp Asn Ala Asn Val Ser Asp Ser
50 55 60
aat aaa gat gga gct tct tat tta att cct caa ggt aaa gaa gct gag 240

Asn	Lys	Asp	Gly	Ala	Ser	Tyr	Leu	Ile	Pro	Gln	Gly	Lys	Glu	Ala	Glu		
65					70					75					80		
tat	aaa	gct	tca	act	gat	ttt	aat	tct	ctt	ttt	acg	aca	act	act	aat	288	
Tyr	Lys	Ala	Ser	Thr	Asp	Phe	Asn	Ser	Leu	Phe	Thr	Thr	Thr	Thr	Asn		
				85					90					95			
gga	ggg	aga	aca	tat	gta	act	aaa	aaa	gat	act	gcg	tca	gca	aat	gag	336	
Gly	Gly	Arg	Thr	Tyr	Val	Thr	Lys	Lys	Asp	Thr	Ala	Ser	Ala	Asn	Glu		
			100					105					110				
att	gcg	aca	tgg	gct	aaa	tct	ata	tca	gct	aat	act	aca	cca	gtt	tcc	384	
Ile	Ala	Thr	Trp	Ala	Lys	Ser	Ile	Ser	Ala	Asn	Thr	Thr	Pro	Val	Ser		
		115					120					125					
act	gtt	act	gag	tca	aat	aat	gat	ggg	act	gag	gtt	att	aat	gtt	tcc	432	
Thr	Val	Thr	Glu	Ser	Asn	Asn	Asp	Gly	Thr	Glu	Val	Ile	Asn	Val	Ser		
	130					135				140							
caa	tat	gga	tat	tat	tat	gtt	tct	agc	act	gtt	aat	aat	gga	gct	gta	480	
Gln	Tyr	Gly	Tyr	Tyr	Tyr	Val	Ser	Ser	Thr	Val	Asn	Asn	Gly	Ala	Val		
145					150				155					160			
att	atg	gtt	aca	tct	gta	act	cca	aat	gct	act	att	cat	gaa	aag	aat	528	
Ile	Met	Val	Thr	Ser	Val	Thr	Pro	Asn	Ala	Thr	Ile	His	Glu	Lys	Asn		
			165				170						175				
act	gat	gcg	aca	tgg	gga	gat	ggg	ggg	gga	aaa	act	gta	gat	caa	aaa	576	
Thr	Asp	Ala	Thr	Trp	Gly	Asp	Gly	Gly	Gly	Lys	Thr	Val	Asp	Gln	Lys		
		180					185					190					
acg	tac	tcg	gtt	ggg	gat	aca	gtc	aaa	tat	act	att	act	tat	aag	aat	624	
Thr	Tyr	Ser	Val	Gly	Asp	Thr	Val	Lys	Tyr	Thr	Ile	Thr	Tyr	Lys	Asn		
		195					200					205					
gca	gtc	aat	tat	cat	ggg	aca	gaa	aaa	gtg	tat	caa	tat	gtt	ata	aag	672	
Ala	Val	Asn	Tyr	His	Gly	Thr	Glu	Lys	Val	Tyr	Gln	Tyr	Val	Ile	Lys		
	210					215				220							
gat	act	atg	cca	tct	gct	tct	gta	gtt	gat	ttg	aac	gaa	ggg	tct	tat	720	
Asp	Thr	Met	Pro	Ser	Ala	Ser	Val	Val	Asp	Leu	Asn	Glu	Gly	Ser	Tyr		
225					230				235					240			
gaa	gta	act	att	act	gat	gga	tca	ggg	aat	att	aca	act	cta	act	caa	768	
Glu	Val	Thr	Ile	Thr	Asp	Gly	Ser	Gly	Asn	Ile	Thr	Thr	Leu	Thr	Gln		
			245					250					255				
ggg	tcg	gaa	aaa	gca	act	ggg	aag	tat	aac	ctg	tta	gag	gaa	aat	aat	816	

Gly Ser Glu Lys Ala Thr Gly Lys Tyr Asn Leu Leu Glu Glu Asn Asn	
260 265 270	
aat ttc acg att act att ccg tgg gca gct acc aat act cca acc gga	864
Asn Phe Thr Ile Thr Ile Pro Trp Ala Ala Thr Asn Thr Pro Thr Gly	
275 280 285	
aat act caa aat gga gct aat gat gac ttt ttt tat aag gga ata aat	912
Asn Thr Gln Asn Gly Ala Asn Asp Asp Phe Phe Tyr Lys Gly Ile Asn	
290 295 300	
aca atc aca gtc act tat aca gga gta tta aag agt gga gct aaa cca	960
Thr Ile Thr Val Thr Tyr Thr Gly Val Leu Lys Ser Gly Ala Lys Pro	
305 310 315 320	
ggt tca gct gat tta cca gaa aat aca aac att gcg acc atc aac ccc	1008
Gly Ser Ala Asp Leu Pro Glu Asn Thr Asn Ile Ala Thr Ile Asn Pro	
325 330 335	
aat act agc aat gat gac cca ggt caa aaa gta aca gtg agg gat ggt	1056
Asn Thr Ser Asn Asp Asp Pro Gly Gln Lys Val Thr Val Arg Asp Gly	
340 345 350	
caa att act ata aaa aaa att gat ggt tcc aca aaa gct tca tta caa	1104
Gln Ile Thr Ile Lys Lys Ile Asp Gly Ser Thr Lys Ala Ser Leu Gln	
355 360 365	
ggt gct ata ttt gtt tta aag aat gct acg ggt caa ttt cta aac ttt	1152
Gly Ala Ile Phe Val Leu Lys Asn Ala Thr Gly Gln Phe Leu Asn Phe	
370 375 380	
aac gat aca aat aac gtt gaa tgg ggc aca gaa gct aat gca aca gaa	1200
Asn Asp Thr Asn Asn Val Glu Trp Gly Thr Glu Ala Asn Ala Thr Glu	
385 390 395 400	
tat aca aca gga gca gat ggt ata att acc att aca ggc ttg aaa gaa	1248
Tyr Thr Thr Gly Ala Asp Gly Ile Ile Thr Ile Thr Gly Leu Lys Glu	
405 410 415	
ggt aca tac tat cta gtt gag aaa aag gct ccc tta ggt tac aat ttg	1296
Gly Thr Tyr Tyr Leu Val Glu Lys Lys Ala Pro Leu Gly Tyr Asn Leu	
420 425 430	
tta gat aac tct cag aag gtt att tta gga gat gga gcc act gat acg	1344
Leu Asp Asn Ser Gln Lys Val Ile Leu Gly Asp Gly Ala Thr Asp Thr	
435 440 445	
act aat tca gat aac ctt tta gtt aac cca act gtt gaa aat aac aaa	1392

Thr Asn Ser Asp Asn Leu Leu Val Asn Pro Thr Val Glu Asn Asn Lys
 450 455 460

ggt act gag ttg cct tca aca ggt ggt att ggt aca aca att ttc tac 1440
 Gly Thr Glu Leu Pro Ser Thr Gly Gly Ile Gly Thr Thr Ile Phe Tyr
 465 470 475 480

att ata ggt gca att tta gta ata gga gca ggt atc gtg ctt gtt gct 1488
 Ile Ile Gly Ala Ile Leu Val Ile Gly Ala Gly Ile Val Leu Val Ala
 485 490 495

cgt cgt cgt tta cgt tct taa 1509
 Arg Arg Arg Leu Arg Ser
 500

<210> 2

<211> 502

<212> PRT

<213> Streptococcus agalactiae

<400> 2

Met Lys Lys Lys Met Ile Gln Ser Leu Leu Val Ala Ser Leu Ala Phe
 1 5 10 15

Gly Met Ala Val Ser Pro Val Thr Pro Ile Ala Phe Ala Ala Glu Thr
 20 25 30

Gly Thr Ile Thr Val Gln Asp Thr Gln Lys Gly Ala Thr Tyr Lys Ala
 35 40 45

Tyr Lys Val Phe Asp Ala Glu Ile Asp Asn Ala Asn Val Ser Asp Ser
 50 55 60

Asn Lys Asp Gly Ala Ser Tyr Leu Ile Pro Gln Gly Lys Glu Ala Glu
 65 70 75 80

Tyr Lys Ala Ser Thr Asp Phe Asn Ser Leu Phe Thr Thr Thr Thr Asn
 85 90 95

Gly Gly Arg Thr Tyr Val Thr Lys Lys Asp Thr Ala Ser Ala Asn Glu
 100 105 110

Ile Ala Thr Trp Ala Lys Ser Ile Ser Ala Asn Thr Thr Pro Val Ser
 115 120 125

Thr Val Thr Glu Ser Asn Asn Asp Gly Thr Glu Val Ile Asn Val Ser
 130 135 140

Gln Tyr Gly Tyr Tyr Tyr Val Ser Ser Thr Val Asn Asn Gly Ala Val
 145 150 155 160

Ile Met Val Thr Ser Val Thr Pro Asn Ala Thr Ile His Glu Lys Asn
 165 170 175

Thr Asp Ala Thr Trp Gly Asp Gly Gly Gly Lys Thr Val Asp Gln Lys
 180 185 190

Thr Tyr Ser Val Gly Asp Thr Val Lys Tyr Thr Ile Thr Tyr Lys Asn
 195 200 205

```

Ala Val Asn Tyr His Gly Thr Glu Lys Val Tyr Gln Tyr Val Ile Lys
 210                215                220
Asp Thr Met Pro Ser Ala Ser Val Val Asp Leu Asn Glu Gly Ser Tyr
225                230                235                240
Glu Val Thr Ile Thr Asp Gly Ser Gly Asn Ile Thr Thr Leu Thr Gln
                245                250                255
Gly Ser Glu Lys Ala Thr Gly Lys Tyr Asn Leu Leu Glu Glu Asn Asn
                260                265                270
Asn Phe Thr Ile Thr Ile Pro Trp Ala Ala Thr Asn Thr Pro Thr Gly
 275                280                285
Asn Thr Gln Asn Gly Ala Asn Asp Asp Phe Phe Tyr Lys Gly Ile Asn
 290                295                300
Thr Ile Thr Val Thr Tyr Thr Gly Val Leu Lys Ser Gly Ala Lys Pro
305                310                315                320
Gly Ser Ala Asp Leu Pro Glu Asn Thr Asn Ile Ala Thr Ile Asn Pro
                325                330                335
Asn Thr Ser Asn Asp Asp Pro Gly Gln Lys Val Thr Val Arg Asp Gly
                340                345                350
Gln Ile Thr Ile Lys Lys Ile Asp Gly Ser Thr Lys Ala Ser Leu Gln
                355                360                365
Gly Ala Ile Phe Val Leu Lys Asn Ala Thr Gly Gln Phe Leu Asn Phe
 370                375                380
Asn Asp Thr Asn Asn Val Glu Trp Gly Thr Glu Ala Asn Ala Thr Glu
385                390                395                400
Tyr Thr Thr Gly Ala Asp Gly Ile Ile Thr Ile Thr Gly Leu Lys Glu
                405                410                415
Gly Thr Tyr Tyr Leu Val Glu Lys Lys Ala Pro Leu Gly Tyr Asn Leu
                420                425                430
Leu Asp Asn Ser Gln Lys Val Ile Leu Gly Asp Gly Ala Thr Asp Thr
                435                440                445
Thr Asn Ser Asp Asn Leu Leu Val Asn Pro Thr Val Glu Asn Asn Lys
                450                455                460
Gly Thr Glu Leu Pro Ser Thr Gly Gly Ile Gly Thr Thr Ile Phe Tyr
465                470                475                480
Ile Ile Gly Ala Ile Leu Val Ile Gly Ala Gly Ile Val Leu Val Ala
                485                490                495
Arg Arg Arg Leu Arg Ser
                500

```

<210> 3

<211> 1692

<212> DNA

<213> Streptococcus agalactiae

<220>

<221> CDS

<222> (1) .. (1692)

<400> 3

```

att tgc att atg gtg atc gta ttc cgg att ata cag ata tta caa ggg      48
Ile Cys Ile Met Val Ile Val Phe Arg Ile Ile Gln Ile Leu Gln Gly
  1               5               10               15

att ata tcc aag atc ctt cag gta cat att att ata agt atg att cac      96
Ile Ile Ser Lys Ile Leu Gln Val His Ile Ile Ile Ser Met Ile His
      20               25               30

gag ata aag atc ccg act caa cta aag atg cct att ata cga cag ata     144
Glu Ile Lys Ile Pro Thr Gln Leu Lys Met Pro Ile Ile Arg Gln Ile
      35               40               45

cta gtc tca tca aat gtt gat aca aca act aag tac aag tac gta aaa     192
Leu Val Ser Ser Asn Val Asp Thr Thr Thr Lys Tyr Lys Tyr Val Lys
      50               55               60

gac gct tac aaa tta gtc ggt tgg tat tat gtt aat cca tat ggt agt     240
Asp Ala Tyr Lys Leu Val Gly Trp Tyr Tyr Val Asn Pro Tyr Gly Ser
      65               70               75               80

att aga cct tat aac ttt tca ggt gct gta act caa gat atc aat tta     288
Ile Arg Pro Tyr Asn Phe Ser Gly Ala Val Thr Gln Asp Ile Asn Leu
      85               90               95

aga gct att tgg cga aag gct gga gat tat cat att ata tac agc aat     336
Arg Ala Ile Trp Arg Lys Ala Gly Asp Tyr His Ile Ile Tyr Ser Asn
      100              105              110

gat gct gtt ggt aca gat gga aag cca gca ttg gat gct tct ggt cag     384
Asp Ala Val Gly Thr Asp Gly Lys Pro Ala Leu Asp Ala Ser Gly Gln
      115              120              125

caa tta caa aca agt aat gag cct act gac cct gat tcc tat gac gat     432
Gln Leu Gln Thr Ser Asn Glu Pro Thr Asp Pro Asp Ser Tyr Asp Asp
      130              135              140

ggc tcc cat tca gcc tta ctg aga cgt ccg aca atg cca gat ggc tat     480
Gly Ser His Ser Ala Leu Leu Arg Arg Pro Thr Met Pro Asp Gly Tyr
      145              150              155              160

cgt ttc cgt ggc tgg tgg tac aat ggt aaa att tat aac cca tat gat     528
Arg Phe Arg Gly Trp Trp Tyr Asn Gly Lys Ile Tyr Asn Pro Tyr Asp
      165              170              175

tcc att gat att gac gcc cat tta gca gat gct aat aaa aat atc acc     576

```

Ser Ile Asp Ile Asp Ala His Leu Ala Asp Ala Asn Lys Asn Ile Thr	
180 185 190	
ata aaa cct gtc att att cca gta gga gat atc aaa tta gaa gat acc	624
Ile Lys Pro Val Ile Ile Pro Val Gly Asp Ile Lys Leu Glu Asp Thr	
195 200 205	
tcc atc aaa tac aat ggt aac ggt ggt act aga gta gaa aat ggt aat	672
Ser Ile Lys Tyr Asn Gly Asn Gly Gly Thr Arg Val Glu Asn Gly Asn	
210 215 220	
gtg gta aca caa gtg gag aca ccg cgt atg gag ttg aat agc aca act	720
Val Val Thr Gln Val Glu Thr Pro Arg Met Glu Leu Asn Ser Thr Thr	
225 230 235 240	
aca att cct gaa aac caa tac ttt aca agg aca ggt tac aac ctt att	768
Thr Ile Pro Glu Asn Gln Tyr Phe Thr Arg Thr Gly Tyr Asn Leu Ile	
245 250 255	
ggg tgg cat cat gat aag gat tta gct gat aca gga cgt gtg gaa ttt	816
Gly Trp His His Asp Lys Asp Leu Ala Asp Thr Gly Arg Val Glu Phe	
260 265 270	
aca gca ggt caa tca ata ggt att gat aac aac ctt gat gca aca aat	864
Thr Ala Gly Gln Ser Ile Gly Ile Asp Asn Asn Leu Asp Ala Thr Asn	
275 280 285	
acc tta tat gct gtt tgg caa cct aaa gaa tac acc gtc gga gta agt	912
Thr Leu Tyr Ala Val Trp Gln Pro Lys Glu Tyr Thr Val Gly Val Ser	
290 295 300	
aaa act gtc gtt gga cta gat gaa gat aag acg aaa gac ttc ttg ttt	960
Lys Thr Val Val Gly Leu Asp Glu Asp Lys Thr Lys Asp Phe Leu Phe	
305 310 315 320	
aat cca agt gaa acg ttg caa caa gag aat ttt ccg ctg aga gat ggt	1008
Asn Pro Ser Glu Thr Leu Gln Gln Glu Asn Phe Pro Leu Arg Asp Gly	
325 330 335	
cag act aag gaa ttt aaa gta cct tat gga act tct ata tca ata gat	1056
Gln Thr Lys Glu Phe Lys Val Pro Tyr Gly Thr Ser Ile Ser Ile Asp	
340 345 350	
gaa caa gcc tac gat gaa ttt aaa gta tct gag tca att aca gaa aaa	1104
Glu Gln Ala Tyr Asp Glu Phe Lys Val Ser Glu Ser Ile Thr Glu Lys	
355 360 365	
aat cta gca act ggt gaa gct gat aaa act tat gat gct acc ggc tta	1152

Asn	Leu	Ala	Thr	Gly	Glu	Ala	Asp	Lys	Thr	Tyr	Asp	Ala	Thr	Gly	Leu		
370						375					380						
caa	tcc	ctg	aca	gtt	tca	gga	gac	gta	gat	att	agc	ttt	acc	aat	aca	1200	
Gln	Ser	Leu	Thr	Val	Ser	Gly	Asp	Val	Asp	Ile	Ser	Phe	Thr	Asn	Thr		
385					390					395					400		
cgt	atc	aag	caa	aaa	gta	cga	cta	cag	aaa	gtt	aat	gtc	gaa	aat	gat	1248	
Arg	Ile	Lys	Gln	Lys	Val	Arg	Leu	Gln	Lys	Val	Asn	Val	Glu	Asn	Asp		
				405					410					415			
aat	aat	ttt	tta	gca	ggg	gca	gtt	ttt	gat	att	tat	gaa	tca	gat	gct	1296	
Asn	Asn	Phe	Leu	Ala	Gly	Ala	Val	Phe	Asp	Ile	Tyr	Glu	Ser	Asp	Ala		
			420					425					430				
aat	ggg	aat	aaa	gct	tca	cat	cct	atg	tat	tca	ggg	ctg	gtg	aca	aac	1344	
Asn	Gly	Asn	Lys	Ala	Ser	His	Pro	Met	Tyr	Ser	Gly	Leu	Val	Thr	Asn		
		435					440					445					
gat	aaa	ggc	ttg	tta	tta	gtg	gat	gct	aat	aac	tac	ctc	agt	ttg	cca	1392	
Asp	Lys	Gly	Leu	Leu	Leu	Val	Asp	Ala	Asn	Asn	Tyr	Leu	Ser	Leu	Pro		
	450					455					460						
gta	gga	aaa	tac	tac	cta	aca	gag	aca	aag	gcc	cct	cca	ggg	tac	cta	1440	
Val	Gly	Lys	Tyr	Tyr	Leu	Thr	Glu	Thr	Lys	Ala	Pro	Pro	Gly	Tyr	Leu		
465					470				475						480		
cta	cct	aaa	aat	gat	gat	ata	tca	gta	tta	gtg	att	tct	acg	gga	gtt	1488	
Leu	Pro	Lys	Asn	Asp	Asp	Ile	Ser	Val	Leu	Val	Ile	Ser	Thr	Gly	Val		
				485					490					495			
acc	ttt	gaa	caa	aat	ggg	aat	aat	gcg	aca	cca	ata	aaa	gag	aat	tta	1536	
Thr	Phe	Glu	Gln	Asn	Gly	Asn	Asn	Ala	Thr	Pro	Ile	Lys	Glu	Asn	Leu		
			500					505					510				
gtg	gat	gga	agt	aca	gta	tat	act	ttt	aaa	att	act	aac	agt	aaa	gga	1584	
Val	Asp	Gly	Ser	Thr	Val	Tyr	Thr	Phe	Lys	Ile	Thr	Asn	Ser	Lys	Gly		
		515					520					525					
aca	gaa	ttg	cct	agt	act	gga	ggg	att	gga	aca	cac	att	tat	atc	cta	1632	
Thr	Glu	Leu	Pro	Ser	Thr	Gly	Gly	Ile	Gly	Thr	His	Ile	Tyr	Ile	Leu		
		530				535					540						
gtt	ggg	tta	gct	tta	gct	cta	cca	tca	gga	tta	ata	tta	tac	tat	cga	1680	
Val	Gly	Leu	Ala	Leu	Ala	Leu	Pro	Ser	Gly	Leu	Ile	Leu	Tyr	Tyr	Arg		
545					550					555					560		
aaa	aaa	ata	tga													1692	

Lys Lys Ile

<210> 4

<211> 563

<212> PRT

<213> Streptococcus agalactiae

<400> 4

```

Ile Cys Ile Met Val Ile Val Phe Arg Ile Ile Gln Ile Leu Gln Gly
 1           5           10           15
Ile Ile Ser Lys Ile Leu Gln Val His Ile Ile Ile Ser Met Ile His
          20           25           30
Glu Ile Lys Ile Pro Thr Gln Leu Lys Met Pro Ile Ile Arg Gln Ile
          35           40           45
Leu Val Ser Ser Asn Val Asp Thr Thr Thr Lys Tyr Lys Tyr Val Lys
          50           55           60
Asp Ala Tyr Lys Leu Val Gly Trp Tyr Tyr Val Asn Pro Tyr Gly Ser
 65           70           75           80
Ile Arg Pro Tyr Asn Phe Ser Gly Ala Val Thr Gln Asp Ile Asn Leu
          85           90           95
Arg Ala Ile Trp Arg Lys Ala Gly Asp Tyr His Ile Ile Tyr Ser Asn
          100          105          110
Asp Ala Val Gly Thr Asp Gly Lys Pro Ala Leu Asp Ala Ser Gly Gln
          115          120          125
Gln Leu Gln Thr Ser Asn Glu Pro Thr Asp Pro Asp Ser Tyr Asp Asp
          130          135          140
Gly Ser His Ser Ala Leu Leu Arg Arg Pro Thr Met Pro Asp Gly Tyr
145           150           155           160
Arg Phe Arg Gly Trp Trp Tyr Asn Gly Lys Ile Tyr Asn Pro Tyr Asp
          165          170          175
Ser Ile Asp Ile Asp Ala His Leu Ala Asp Ala Asn Lys Asn Ile Thr
          180          185          190
Ile Lys Pro Val Ile Ile Pro Val Gly Asp Ile Lys Leu Glu Asp Thr
          195          200          205
Ser Ile Lys Tyr Asn Gly Asn Gly Gly Thr Arg Val Glu Asn Gly Asn
          210          215          220
Val Val Thr Gln Val Glu Thr Pro Arg Met Glu Leu Asn Ser Thr Thr
225           230           235           240
Thr Ile Pro Glu Asn Gln Tyr Phe Thr Arg Thr Gly Tyr Asn Leu Ile
          245          250          255
Gly Trp His His Asp Lys Asp Leu Ala Asp Thr Gly Arg Val Glu Phe
          260          265          270
Thr Ala Gly Gln Ser Ile Gly Ile Asp Asn Asn Leu Asp Ala Thr Asn
          275          280          285
Thr Leu Tyr Ala Val Trp Gln Pro Lys Glu Tyr Thr Val Gly Val Ser
          290          295          300
Lys Thr Val Val Gly Leu Asp Glu Asp Lys Thr Lys Asp Phe Leu Phe

```

305		310		315		320
Asn Pro Ser Glu Thr Leu Gln Gln Glu Asn Phe Pro Leu Arg Asp Gly						
	325		330		335	
Gln Thr Lys Glu Phe Lys Val Pro Tyr Gly Thr Ser Ile Ser Ile Asp						
	340		345		350	
Glu Gln Ala Tyr Asp Glu Phe Lys Val Ser Glu Ser Ile Thr Glu Lys						
	355		360		365	
Asn Leu Ala Thr Gly Glu Ala Asp Lys Thr Tyr Asp Ala Thr Gly Leu						
	370		375		380	
Gln Ser Leu Thr Val Ser Gly Asp Val Asp Ile Ser Phe Thr Asn Thr						
385		390		395		400
Arg Ile Lys Gln Lys Val Arg Leu Gln Lys Val Asn Val Glu Asn Asp						
	405		410		415	
Asn Asn Phe Leu Ala Gly Ala Val Phe Asp Ile Tyr Glu Ser Asp Ala						
	420		425		430	
Asn Gly Asn Lys Ala Ser His Pro Met Tyr Ser Gly Leu Val Thr Asn						
	435		440		445	
Asp Lys Gly Leu Leu Leu Val Asp Ala Asn Asn Tyr Leu Ser Leu Pro						
	450		455		460	
Val Gly Lys Tyr Tyr Leu Thr Glu Thr Lys Ala Pro Pro Gly Tyr Leu						
465		470		475		480
Leu Pro Lys Asn Asp Asp Ile Ser Val Leu Val Ile Ser Thr Gly Val						
	485		490		495	
Thr Phe Glu Gln Asn Gly Asn Asn Ala Thr Pro Ile Lys Glu Asn Leu						
	500		505		510	
Val Asp Gly Ser Thr Val Tyr Thr Phe Lys Ile Thr Asn Ser Lys Gly						
	515		520		525	
Thr Glu Leu Pro Ser Thr Gly Gly Ile Gly Thr His Ile Tyr Ile Leu						
	530		535		540	
Val Gly Leu Ala Leu Ala Leu Pro Ser Gly Leu Ile Leu Tyr Tyr Arg						
545		550		555		560
Lys Lys Ile						

<210> 5

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic
oligonucleotide

<400> 5

ctaggtggat ccttcggcaa t

21

<210> 6
<211> 10
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Synthetic
oligonucleotide

<400> 6
cgattgccga

10

<210> 7
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Synthetic
oligonucleotide

<400> 7
aggcaactgt gctaaccgag ggaat

25

<210> 8
<211> 11
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Synthetic
oligonucleotide

<400> 8
cgattccctc g

11

<210> 9
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Consensus
sequence

<400> 9

Leu Pro Xaa Thr Gly

1 5

<210> 10

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Consensus
sequence

<400> 10

Leu Pro Ser Thr Gly Gly

1 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/17082**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.7; 530/350, 825; 424/244.1, 184.1, 236.1, 234.1, 193.1; 435/320.1, 69.3, 71.1, 7.1, 7.34

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

DIALOG, MEDLINE, BIOSIS, EMBASE, WEST

spb1, spb2, GBS, SEQ ID NOs: 1, 2, 3 and 4, streptococc?, inventors' names

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P ----- Y,P	ADDERSON et al. Genetic subtraction identifies a novel respiratory epithelial adhesin/invasin of type III group B streptococcus (GBS). Pediatr. Res. 30 March 2000, Vol. 47, No. 4, page 337A, abstract 1992, see entire abstract.	1-16, 56 ----- 17-32, 33-55, 57-67

☐ Further documents are listed in the continuation of Box C.
 ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 31 OCTOBER 2000	Date of mailing of the international search report 14 NOV 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer S. DEVI, Ph.D. <i>Jaige Budgers</i> Telephone No. (703) 308-1235

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/17082

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☒
☐

The additional search fees were accompanied by the applicant's protest.
No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/17082

A. CLASSIFICATION OF SUBJECT MATTER: IPC (7):

C07H 21/04; C07K 1/00, A61K 39/09, 39/38, 39/39, 38/385; C12N 15/00, 15/09; C12P 21/04, G01N 33/53, 33/569

A. CLASSIFICATION OF SUBJECT MATTER: US CL :

536/23.7; 530/350, 825; 424/244.1, 184.1, 236.1, 234.1, 193.1; 435/320.1, 69.3, 71.1, 7.1, 7.34

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

Group I, claim(s) 1-16, 33-35 and 56-60, drawn to an isolated nucleic acid molecule encoding SEQ ID NO: 2, a vector and a host cell comprising the same and a method of immunizing a mammal by administering a protein comprising SEQ ID NO: 2.

Group II, claim(s) 17-32, 38-40, 61-64 and 67, drawn to an isolated nucleic acid molecule encoding SEQ ID NO: 4, a vector and a host cell comprising the same and a method of immunizing a mammal by administering a protein comprising SEQ ID NO: 4.

Group III, claim(s) 42, 50 and 54, drawn to a diagnostic method comprising analyzing for the presence or absence of SEQ ID NO: 2.

Group IV, claims 43, 51 and 55, drawn to a diagnostic method comprising analyzing for the presence or absence of SEQ ID NO: 4.

Claims 36, 37, 65 and 66 are considered linking claims and would be joined with Group I or II, if elected.

Claims 41, 44-49, 52 and 53 are considered linking claims and would be joined with Group II or III, if elected.

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Group I is directed to a nucleic acid molecule of SEQ ID NO:1, a protein with amino acid sequence of SEQ ID NO: 2, a vector and a host cell comprising the nucleic acid molecule and the first method of use of the protein in immunization. This is a permitted category under PCT Rule 13.2. Group II is directed to a second product, a nucleic acid molecule of SEQ ID NO:3, protein with amino acid sequence of SEQ ID NO: 4, a vector and a host cell comprising the nucleic acid molecule and a method of immunizing a mammal by administering the protein. Groups III and IV are directed respectively to a second method of use of the gene products, i.e., diagnostic methods for analyzing spb1 and spb2 gene products respectively. The SEQ ID NO.s of inventions I and II are the special technical features, which distinct from one another in their structure and specificity. The inventions therefore lack unity because the special technical feature is not a unifying feature.

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
28 December 2000 (28.12.2000)

PCT

(10) International Publication Number
WO 00/78787 A1

(51) International Patent Classification⁷: C07H 21/04,
C07K 1/00, A61K 39/09, 39/38, 39/39, 38/385, C12N
15/00, 15/09, C12P 21/04, G01N 33/53, 33/569

(21) International Application Number: PCT/US00/17082

(22) International Filing Date: 21 June 2000 (21.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/140,084 21 June 1999 (21.06.1999) US

(71) Applicant (for all designated States except US): UNIVER-
SITY OF UTAH RESEARCH FOUNDATION [US/US];
Suite 110, 615 Arapen Drive, Salt Lake City, UT 84108
(US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ADDERSON, Elis-
abeth [US/US]; St. Jude Children's Research Hospital,
Room D 2038, 332 N. Lauderdale, Memphis, TN 38105
(US). BOHNSACK, John [US/US]; University of Utah
Health Sciences Center, Dept. of Pediatrics, 50 North Med-
ical Drive, Salt Lake City, UT 84132 (US).

(74) Agent: MADSON & METCALF; 15 West South Temple,
Suite 900, Salt Lake City, UT 84101 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— With international search report.

(48) Date of publication of this corrected version:

5 April 2001

(15) Information about Correction:

see PCT Gazette No. 14/2001 of 5 April 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: ISOLATED GENES FROM VIRULENT GROUP B *STREPTOCOCCUS AGALACTIAE*

(57) Abstract: The present invention relates to the identification and prevention of infections by virulent forms of Group B strepto-
cocci. Disclosed herein is the identification of two genes, *spb1* and *spb2*, that are specific to virulent type III-3 GBS. Also disclosed
herein are diagnostic methods for detecting virulent GBS infections and methods of immunizing a mammal against these bacteria.

WO 00/78787 A1

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
28 December 2000 (28.12.2000)

PCT

(10) International Publication Number
WO 00/78787 A1

(51) International Patent Classification⁷: **C07H 21/04**,
C07K 1/00, A61K 39/09, 39/38, 39/39, 38/385, C12N
15/00, 15/09, C12P 21/04, G01N 33/53, 33/569

(21) International Application Number: PCT/US00/17082

(22) International Filing Date: 21 June 2000 (21.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/140,084 21 June 1999 (21.06.1999) US

(71) Applicant (for all designated States except US): **UNIVERSITY OF UTAH RESEARCH FOUNDATION** [US/US];
Suite 110, 615 Arapleen Drive, Salt Lake City, UT 84108
(US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ADDERSON, Elisabeth** [US/US]; St. Jude Children's Research Hospital,
Room D 2038, 332 N. Lauderdale, Memphis, TN 38105
(US). **BOHNSACK, John** [US/US]; University of Utah
Health Sciences Center, Dept. of Pediatrics, 50 North Medical
Drive, Salt Lake City, UT 84132 (US).

(74) Agent: **MADSON & METCALF**; 15 West South Temple,
Suite 900, Salt Lake City, UT 84101 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(48) Date of publication of this corrected version:

27 June 2002

(15) Information about Corrections:

see PCT Gazette No. 26/2002 of 27 June 2002, Section II

Previous Correction:

see PCT Gazette No. 14/2001 of 5 April 2001, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ISOLATED GENES FROM VIRULENT GROUP B *STREPTOCOCCUS AGALACTIAE*

(57) Abstract: The present invention relates to the identification and prevention of infections by virulent forms of Group B streptococci. Disclosed herein is the identification of two genes, *spb1* and *spb2*, that are specific to virulent type III-3 GBS. Also disclosed herein are diagnostic methods for detecting virulent GBS infections and methods of immunizing a mammal against these bacteria.

WO 00/78787 A1

ISOLATED GENES FROM VIRULENT GROUP B *STREPTOCOCCUS* *AGALACTIAE*

1. FIELD OF THE INVENTION

The present invention relates to genes isolated from Group B streptococci (“GBS”). More specifically, the present invention relates to genes that are specific to virulent forms of GBS and methods of using such genes and their products for the diagnosis and treatment of GBS infections.

2. TECHNICAL BACKGROUND

Group B streptococci (“GBS”) are a common cause of disease in newborns, pregnant women, and other persons. Common manifestations of these infections include bacteremia, pneumonia, meningitis, endocarditis, and osteoarticular infections. C.J. Baker & M.S. Edwards, *Group B Streptococcal Infections*, in *Infectious Disease of the Fetus and Newborn Infant*, 980-1054 (J.S. Remington & J.O. Klein, eds., 1995); P. Munoz et al., *Arch Int Med* **157**:213-216 (1997).

Nearly 3 out of every 1,000 children born are infected with an invasive form of GBS disease. While GBS disease is of great concern in neonates, GBS is also an important pathogen in the general population, in which the incidence of invasive GBS disease is nearly 8 in 100,000. Of these infections, the mortality rate can be as high as 30%.

During childbirth, GBS can pass from the mother to the newborn. By one estimate, up to 30% of pregnant women carry GBS at least temporarily in the vagina or rectum without symptoms. Infants born to these women become colonized with GBS during delivery. Baker & Edwards, *supra*. Aspiration of infected amniotic fluid or vaginal secretions allow GBS to gain access to the lungs. Adhesion to, and invasion of, respiratory epithelium and endothelium appear to be critical factors in early onset neonatal infection. Baker & Edwards, *supra*; C.E. Rubens et al., *J Inf Dis* **164**:320-330 (1991). Subsequent steps in infection, such as blood stream invasion and the establishment of metastatic local infections have not been clarified. The pathogenesis of neonatal infection occurring after the first week of life is also not well understood. Gastrointestinal colonization may be more important than a respiratory focus in late onset neonatal disease. Baker & Edwards, *supra*. Considerable evidence suggests that invasion of brain microvascular endothelial cells by GBS is the initial step in the pathogenesis of

meningitis. GBS are able to invade human brain microvascular endothelial cells and type III GBS, which are responsible for the majority of meningitis, accomplish this 2-6 times more efficiently than other serotypes. V. Nizet et al. *Infect Immun* **65**:5074-5081 (1997).

Because GBS is widely distributed among the population and is an important pathogen in newborns, pregnant women are commonly tested for GBS at 26 to 28 weeks of pregnancy. Much of GBS neonatal disease is preventable by administration of prophylactic antibiotics during labor to women who test positive or display known risk factors. However, these antibiotics programs do not prevent all GBS disease. The programs are deficient for a number of reasons. First, the programs can be inefficient. Second, it is difficult to ensure that all healthcare providers and patients comply with the testing and treatment. And finally, if new serotypes or antibiotic resistance emerges, the antibiotic programs may fail altogether. Currently available tests for GBS are inefficient. These tests may provide false negatives. Furthermore, the tests are not specific to virulent strains of GBS. Thus, antibiotic treatment may be given unnecessarily and add to the problem of antibiotic resistance. Although a vaccine would be advantageous, none are yet commercially available.

Traditionally, GBS have been divided into 9 serotypes according to the immunologic reactivity of the polysaccharide capsule. H.M. Blumberg et al., *J Inf Dis* **173**: 365-373 (1996). Serotype III GBS cause 60-70% of all infections and almost all meningitis. Baker & Edwards, *supra*. Type III GBS can be subdivided into three groups of related strains based on the analysis of restriction digest patterns (RDPs) produced by digestion of chromosomal DNA with *Hind* III and *Sse*8387 I. Y. Nagano et al., *J Med Micro* **35**:297-303 (1991); S. Takahashi et al., *J Inf Dis* **177**:1116-1119 (1998). Figure 1 illustrates a comparison of *Hind* III and *Sse*8387 I RDP typing of 62 type III isolates from Salt Lake City, Utah and Tokyo, Japan. Isolates were classified into types based on the similarity of the restriction digest patterns produced by *Hind* III or *Sse*8387 I digestion of chromosomal DNA. The two methods divided the isolates into RDP types containing exactly the same isolates: III-3 contains isolates 1-41, II-2 contains isolates 42-59, and II-1 contains isolates 60-62.

Over 90% of invasive type III GBS neonatal disease in Tokyo, Japan and in Salt Lake City, Utah is caused by bacteria from one of three RDP types, termed RDP type III-3, while RDP type III-2 are significantly more likely to be isolated from vagina than from

blood or CSF. These results suggest that this genetically-related cluster of type III-3 GBS are more virulent than III-2 strains and could be responsible for the majority of invasive type III disease globally.

From the foregoing, it will be appreciated that it would be a significant advancement in the art to provide one or more markers that are specific to virulent type III-3 GBS. It would be a further advancement to provide a method to exploit these markers for clinical identification of virulent type III-3 GBS. It would be a further advancement to provide methods for producing vaccines against type III-3 GBS.

Such compositions and methods are disclosed herein.

3. BRIEF SUMMARY OF THE INVENTION

The present invention relates to markers specific to type III-3 GBS. These markers, the *spb1* and *spb2* gene products (SEQ ID NO: 2 and SEQ ID NO: 4, respectively), are encoded by the *spb1* (SEQ ID NO: 1) and *spb2* (SEQ ID NO: 3) genes. The invention also provides these genes and gene products in substantially purified form.

In certain other embodiments, the present invention relates to recombinant vectors which incorporate the *spb1* gene or other nucleic acid molecules that code for the *spb1* gene product. The recombinant vector may be a plasmid. In certain embodiments, the recombinant vector is a prokaryotic or eukaryotic expression vector. In certain preferred embodiments, the nucleic acid molecule is operably linked to a heterologous promoter and/or other expression control elements, such as heterologous enhancers and polyadenylation sequences.

In certain other embodiments, the present invention relates to recombinant vectors which incorporate the *spb2* gene or other nucleic acid molecules that code for the *spb2* gene product. The recombinant vector may be a plasmid. In certain embodiments, the recombinant vector is a prokaryotic or eukaryotic expression vector. In certain preferred embodiments, the nucleic acid molecule is operably linked to a heterologous promoter and/or other expression control elements.

The present invention also provides host cells comprising the *spb1* and/or *spb2* genes. In other embodiments, a host cell of the present invention comprises nucleic acid molecules that code for the *spb1* and/or *spb2* gene products. The host cell may be a prokaryotic or eukaryotic host cell.

The present invention also relates to diagnostic methods for determining whether a mammal is infected or colonized by virulent GBS. In certain embodiments, a diagnostic method comprises the steps of (1) collecting a bodily fluid or culture from the mammal and (2) analyzing the bodily fluid or culture for the presence or absence of one or more gene products specific to type III-3 GBS, wherein the presence of one or more gene products specific to type III-3 GBS indicates infection or colonization by virulent GBS. The mammal may be a human. Alternatively, the mammal may be a laboratory, domestic, or agricultural animal. The bodily fluid or culture may be any bodily fluid or culture that is typically analyzed for the presence of bacteria. For example, the bodily fluid or culture may be a vaginal or rectovaginal culture. The bodily fluid or culture may also be a throat culture. The bodily fluid or culture may also be an endotracheal tube aspirant, fluid from a bronchoalveolar lavage, or tissue from a lung biopsy. In certain embodiments, the bodily fluid or culture is blood, serum, amniotic fluid, cerebrospinal fluid, or joint fluid. Other sources of material will be apparent to those of skill in the art. In certain embodiments, a diagnostic method of the present invention comprises analyzing a sample for the presence or absence of the *spb1* and/or *spb2* gene product(s).

In certain embodiments of a diagnostic method of the present invention, the polymerase chain reaction ("PCR") is used to identify the presence or absence of the *spb1* and/or *spb2* gene(s). In certain other embodiments, antibodies are used to identify the presence or absence of the *spb1* and/or *spb2* gene products. The antibodies may be monoclonal or polyclonal antibodies.

The present invention also relates to GBS vaccines. In certain embodiments, the present invention provides vaccines comprising the *spb1* gene product, i.e., a protein comprising the amino acid sequence of SEQ ID NO: 2. In certain other embodiments, a vaccine comprises the *spb2* gene product, i.e., a protein comprising the amino acid sequence of SEQ ID NO: 4. In certain preferred embodiments, a vaccine comprises both the *spb1* and *spb2* gene products. The vaccine may include an adjuvant, such as alum. In certain other embodiments, the *spb1* and/or *spb2* gene(s) may be introduced into a mammal using either naked DNA or other gene therapy techniques to induce an immune response against type III GBS.

The present invention further provides methods of immunizing a mammal against GBS infection. In certain embodiments, such methods comprise administering to the

mammal a vaccine comprising an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 2. In certain other embodiments, a method of the present invention comprises administering to the mammal a vaccine comprising an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 4. The vaccine may also contain a mixture of the *spb1* and *spb2* gene products. Vaccines used in the methods of the present invention may further comprise an adjuvant, such as alum.

These and other features and advantages of the present invention will become more fully apparent from the following detailed description.

4. SUMMARY OF THE DRAWINGS

Figure 1 illustrates a comparison of *HindIII* and *Sse83871* RDP typing of 62 type III GBS isolates from Salt Lake City, Utah and Tokyo, Japan. Isolates were classified into types based on the similarity of the restriction digest patterns produced by *HindIII* or *Sse83871* digestion of chromosomal DNA. The two methods divided isolates into RDP types containing exactly the same isolates: III-3 contains isolates 1 - 41, III-2 contains isolates 42 - 59, and III-1 contains isolates 60 - 62.

Figure 2 illustrates a dot blot hybridization of probe 1 with genomic DNA isolated from type III GBS. 10 µg of genomic DNA from each of 62 type III GBS strains was transferred to nylon membrane. Radiolabeled probe 1 hybridized with DNA from all III-3 strains (rows A - D) including the original type III-3 strain (well E1). The probe failed to hybridize with DNA from III-2 strains (F1 - F10, G1 - G7) including the original strain used in the subtraction hybridization (well E10) and III-1 strains (wells H1 - H3). The same pattern of hybridization was observed using clone 3 and 11 probes.

5. DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to the identification and prevention of infections by virulent forms of GBS. The present invention also relates to isolated genes specific to type III-3 GBS. These genes, *spb1* and *spb2*, encode the *spb1* and *spb2* gene products.

The *spb1* and *spb2* genes or other nucleic acid molecules coding for the *spb1* or *spb2* gene products may be incorporated into a recombinant vector using methods known in the art. See, e.g., 1-3 J. Sambrook et al., *Molecular Cloning: A Laboratory Manual*

(2d ed. 1989). Recombinant vectors include any genetic element, such as a plasmid, phage, transposon, cosmid, chromosome, virus, etc., that is capable of replication when associated with the proper control elements and that can transfer gene sequences between cells. Thus, the term includes cloning and expression vehicles.

A nucleic acid molecule of the present invention may be operably linked to expression control sequences, such as heterologous promoters. Examples include, but are not limited to, viral promoters such as the SV40 early promoter and the CMV immediate early promoter region, bacterial promoters, mammalian promoters, inducible promoters, synthetic promoters, hybrid promoters, and the like. Other expression control sequences are known in the art and include polyadenylation signals, transcription termination sequences, upstream regulatory domains, origins of replication, internal ribosome entry sites ("IRES"), and enhancers. These expression control sequences collectively provide for the replication, transcription and translation of a coding sequence in a recipient cell. Not all of these control sequences need always be present in a recombinant vector, so long as the selected coding sequence is capable of being replicated, transcribed and translated in an appropriate host cell.

Recombinant vectors can be constructed to include selectable markers. Suitable markers include genes which confer antibiotic resistance or sensitivity, or impart color, or change the antigenic characteristics when host cells which have been transfected with the recombinant vectors are grown in an appropriate selective medium. Suitable markers are known to those of skill in the art.

The discovery of type III-3 GBS-specific gene products will allow clinicians to diagnose and treat infection and colonization with virulent GBS. For example, hybridization-based assays may be used to determine whether a GBS isolate is type III-3. Figure 2 illustrates the results of hybridization assays with a III-3-specific probe. Likewise, PCR may be used to detect the presence or absence of either the *spb1* gene or the *spb2* gene (or both) in samples from patients. PCR methods are described generally in C.R. Newton & A. Graham, *PCR* (2nd. ed. 1997); *PCR: Essential Techniques* (J.F. Burke ed., 1996). Patients who are infected with type III-3 GBS may then receive appropriate antibiotic therapy.

Antibodies may also be used to detect the presence or absence of the *spb1* and/or *spb2* gene product(s). Methods for preparing both monoclonal and polyclonal antibodies are described in, e.g., E. Harlow & D. Lane, *Antibodies: A Laboratory Manual* (1988).

The present invention also relates to methods for producing type III-3 GBS vaccines. See generally *Vaccine Protocols* (A. Robinson, G.H. Farrar & C.N. Wiblin eds. 1996). In certain embodiments, the *spb1* and/or *spb2* gene product(s) may be used to immunize against GBS. These gene products may be produced in large quantities using techniques that are known in the art. For example, the appropriate gene or genes may be linked to a prokaryotic promoter and expressed in bacteria. The gene products may then be purified using conventional techniques and used to vaccinate at-risk individuals. Alternatively, the appropriate gene or genes may be linked to a eukaryotic promoter and enhancer (e.g., yeast, baculovirus, SV40, etc.) and expressed in an appropriate cell type. The gene products may then be purified using conventional techniques.

The *spb1* and/or *spb2* gene products, or immunogenic fragments thereof, may stimulate an immune response when administered to a host. Recombinantly produced proteins are especially desirable, as they can be produced in large amounts and purified. Furthermore, recombinantly produced proteins may be engineered to maximize desirable activities and to minimize unwanted effects.

The recombinantly produced *spb1* and/or *spb2* gene products may be used as carrier proteins for a polysaccharide-protein or oligosaccharide-protein conjugate vaccine. See, e.g., R. Schneerson, et al., *Infect Immun* **60**:3528-3532 (1992) (describing a *Pneumococcus-pertussis* toxin conjugate vaccine). For example, *Haemophilus influenzae* B vaccines have been produced by conjugating a tetanus toxoid; a *Corynebacterium* toxin, CRM₁₉₇ (which is a mutant diphtheria toxin); and a *Neisseria* outer membrane protein. Oligo- and polysaccharides from GBS might be used in a vaccine. Oligosaccharide- and polysaccharide-protein conjugates alter the immunological properties of the polysaccharide or oligosaccharide and may improve the immune response.

An adjuvant may be used to enhance the immune response to a vaccine containing the *spb1* and/or *spb2* gene products. An adjuvant is any substance that enhances the immune response to an antigen. Without being bound by any particular theory, adjuvants may act by retaining the antigen locally near the site of administration to produce a depot

effect, facilitating the slow, sustained release of the antigen to cells of the immune system. Adjuvants may also attract cells of the immune system. Aluminum hydroxide and aluminum phosphate (collectively and commonly referred to as “alum”) are routinely used as adjuvants in human and veterinary vaccines. Currently, alum is the only adjuvant licensed for human use, although a number of experimental adjuvants are being tested.

5 The *spb1* and/or *spb2* gene(s) may also be introduced into a mammal using either naked DNA or other gene therapy techniques to induce an immune response against virulent GBS.

 All publications, patents, and patent applications cited in this application are hereby incorporated by reference. U.S. Patent Application Serial No. 60/140,084 is
10 hereby incorporated by reference in its entirety.

6. EXAMPLES

 The following examples are given to illustrate several embodiments which have been made within the scope of the present invention. It is to be understood that these
15 examples are neither comprehensive nor exhaustive of the many types of embodiments which can be prepared in accordance with the present invention.

Example 1 - Isolation of the *spb1* and *spb2* genes

 Bacterial factors that contribute to the increased virulence of III-3 strains can be identified by characterizing the differences between the genetic composition of III-3 and
20 III-2 strains. Such genetic differences will be found in the bacterial chromosomes since these strains do not contain plasmids. Takahashi et al., *supra*.

 To identify genes present in virulent type III-3 GBS strains and not in the avirulent type III-2 strains, a modification of the technique described by Lisitsyn et al., *Science* **259**:946-951 (1993), was used. High molecular weight genomic DNA from an
25 invasive RDP type III-3 GBS strain (strain 874391) and a colonizing (“avirulent”) RDP type III-2 strain (strain 865043) were prepared by cell lysis with mutanolysin and Proteinase K digestion. Y. Nagano et al., *supra*. For genetic subtraction, genomic DNA from both strains was digested with *Taq* I. *Taq* I-digested DNA from the virulent strain was mixed with two complementary oligonucleotides, *Taq*A (5'-CTAGGTGGA-
30 TCCTTCGGCAAT-3' (SEQ ID NO: 5)) and *Taq*B (5'-CGATTGCCGA-3' (SEQ ID NO: 6)), heated to 50°C for 5 minutes, then allowed to cool slowly to 16°C in T4 ligase

buffer. Oligonucleotides were ligated to the virulent strain DNA by incubation with 20 units of T4 ligase at 16°C for 12 hours. After ligation, 500 ng of DNA from the virulent strain, with ligated linkers, and 40 µg of DNA from the avirulent strain, without linkers, was mixed together, denatured by heating, and hybridized at 68°C for 20 hours.

Ten percent of the resulting hybridization mixture was incubated with *Taq* DNA polymerase and dNTPs to fill in the ends of annealed virulent strain DNA. The hybridized DNA was amplified by *Taq* DNA polymerase for 10 cycles using the *TaqA* oligonucleotide as the forward and reverse amplification primer. After amplification, single stranded products remaining after amplification were digested with mung bean nuclease. Twenty percent of the resulting product was then reamplified for 20 cycles. This process of subtraction followed by PCR amplification results in enhanced amplification of DNA segments from the III-3 strains that do not hybridize with DNA segments from the III-2 strains.

A total of four cycles of subtraction and amplification were carried out using successively smaller quantities of III-3 specific PCR products. Two pairs of oligonucleotides were used for subtraction, which were alternated with successive rounds of subtraction-amplification. The oligos were *TaqA* (SEQ ID NO: 5) and *TaqB* (SEQ ID NO: 6) (the first pair) and *TaqE* (5'-AGGCAACTGTGCTAACCGAGGGAAT-3' (SEQ ID NO: 7)) and *TaqF* (5'-CGATTCCCTCG-3' (SEQ ID NO: 8)) (the second pair). The final amplification products were ligated into pBS KS+ vector and transformed into competent XL1-Blue strain *E. coli*. Thirteen clones were randomly selected for analysis. Cross hybridization studies of the 13 inserts revealed that 6 were unique. These probes were used in slot and dot blot experiments to determine whether subtraction was successful and to identify probes hybridizing with all III-3 strains. Each of the 6 unique probes hybridized with the parental III-3 virulent strain, while none of the probes hybridized with the avirulent III-2 strains. Three of the amplified sequence tags (clones 1, 3 and 11) hybridized with genomic DNA from all 62 type III-3 isolates, but did not hybridize with DNA prepared from the III-2 and III-1 isolates.

Figure 2 illustrates a dot blot hybridization of type III GBS genomic DNA hybridized with a clone 1 probe. 10 µg of genomic DNA from each of 62 type III GBS strains was transferred to nylon membrane. Radiolabeled clone 1 probe hybridized with DNA from all III-3 strains (rows A-D) including the original type III-3 strain (well E-1).

The probe failed to hybridize with DNA from III-2 strains (F1-F10, G1-G7) including the original strain used in the subtraction hybridization (well E10) and III-1 strains (wells H1-H3). The same pattern of hybridization was observed using clone 3 and 11 probes. These data demonstrate the feasibility of identifying genes unique to III-3 strains by this method of PCR-based subtractive hybridization, and further support the validity of the RDP typing for identifying genetically-related type III GBS.

The three GBS type III-3-specific sequence tags are short (130-360 bp). To obtain additional sequence information, a genomic GBS III-3 library was constructed. High molecular weight GBS genomic DNA was partially digested with *Bgl* II and cloned into λ FIX II phage vector. Phage were packaged and the library, consisting of 1.7×10^5 recombinant phage containing inserts with a mean size of about 18 kb (totaling approximately 3×10^9 bp), was amplified once. Multiple plaques hybridizing with each of the III-3 GBS-specific probes were purified for further characterization.

Three overlapping genomic clones hybridizing with probe 1 were identified, with approximate sizes of 9, 22, and 23 kb. Since the boundaries of GBS III-3 specific DNA are not known, smaller fragments were subcloned and the DNA was verified present in virulent GBS strains before proceeding with further characterization. The first segment examined is a 6.4 kb *Sal* I-*Bgl* II fragment. This genomic DNA is present in all RDP type III-3 strains and in none of 17 RDP type III-2 strains.

Over 90% of this genomic DNA fragment has been sequenced and found to contain 5 open reading frames ("ORFs"). Three of these are likely to be authentic genes. They feature ATG start sites, are preceded by potential ribosomal binding sites and, in two cases, stop codons are followed by palindromic sequences that may represent transcriptional terminators. They are relatively short, however, and do not show significant homology at the nucleic acid or amino acid level with sequences registered with Genbank or the available bacterial genome databases.

The two other ORFs appear to be more obvious candidates for virulence genes. The *spb1* gene includes a 1509 bp ORF that is preceded by a potential ribosomal binding site 10 bases upstream from an ATG start codon. The predicted protein (502 amino acids and Mr 53,446) has the characteristics of a cell wall-bound protein. The N-terminus of the predicted protein is a hydrophilic, basic stretch of 6 amino acids followed by a 23 amino acid hydrophobic, proline-rich core, consistent with a signal peptide. The

hydrophilic mature protein terminates in a typical LPXTG (SEQ ID NO: 9) domain that immediately precedes a hydrophobic 20 amino acid core and a short, basic hydrophilic terminus.

The nucleotide sequence is not homologous to sequences of other known bacterial genes. The translated amino acid sequence, however, shares segmental homology with a number of characterized proteins, including the fimbrial type 2 protein of *Actinomyces naeslundii* (27% identity over 350 amino acids) and the fimbrial type 1 protein of *Actinomyces viscosus* (25% homology over 420 amino acids), the T6 surface protein of *S. pyogenes* (23% identity over 359 amino acids), and the *hsf* (27% identity over 260 amino acids) and HMW1 adhesins (25% identity over 285 amino acids) of *Haemophilus influenzae*. See M.K. Yeung & J.O. Cisar, *J Bacteriol* **172**:242-2468 (1990); O. Schneewind, et al., *J Bacteriol* **172**:3310-3317 (1990); J.W. St. Geme III, et al., *J Bacteriol* **178**:6281-6287 (1996); J.W. St. Geme III, *Infect Immun* **62**: 3881-3889 (1994). The function of the *S. pyogenes* T6 protein is unknown. Each of the other homologues plays a role in bacterial adhesion and/or invasion.

The second ORF, *spb2*, terminates 37 bp upstream from *spb1* and is in the same transcriptional orientation. This 1692 bp ORF has a deduced amino acid sequence of 563 residues and Mr 64,492. It shares 50.5% nucleic acid identity and 20.7% amino acid identity with *spb1*. Conservation is highest in the carboxy-terminal regions, including a shared LPSTGG (SEQ ID NO: 10) motif. In contrast to *spb1*, *spb2* does not have an obvious signal sequence. Its secretion may be mediated by carboxy-terminal recognition sequences or by accessory peptides. T. Michiels, et al., *Infect Immun* **58**:2840-2849 (1990). The deduced amino acid sequence of *spb2* is also homologous with *S. pyogenes* T6 and *Actinomyces naeslundii* proteins, and to *Listeria monocytogenes* internalin A (22% identity over 308 amino acids)—again, proteins important in adhesion and invasion.

Neither of the predicted gene products has the repetitive structure of previously described GBS surface proteins such as the C and Rib proteins or of *L. monocytogenes* internalin family members. L.C. Madoff et al., *Infect Immun* **59**:2638-2644 (1991); J. Gaillard, et al., *Cell* **65**: 1127-1141 (1991). Hybridization of the originating strain 874391 genomic DNA with an *spb1* probe under low stringency conditions identifies a single band in *EcoR* I, *Sal* I and *Sst* I digests (data not shown), suggesting that a single

copy of *spb1* is present in this strain and that *spb1* is not a member of a significantly homologous “family” of genes.

Example 2 - Adherence and Invasion Assays Using *spb1*⁻ Bacteria

Genomic subtraction identified a 1509 bp ORF, *spb1*, which is present in virulent RDP type III-3 GBS and not in RDP type III-2 strains. The predicted 53 kD protein product has the characteristics of a typical gram positive cell-wall bound protein. The nucleic acid sequence of *spb1* is not homologous to sequences of other known bacterial genes, however the translated amino acid sequence shares segmental homology with several characterized adhesins, including *Actinomyces* fimbrial proteins and *H. influenzae* HMW1, suggesting that Spb1 might contribute to GBS adhesion or invasion. A *spb1*⁻ isogenic deletion mutant GBS strain was created by homologous recombination and the ability of the *spb1*⁻ mutant to adhere to and invade A549 respiratory epithelial cells was determined. Compared to the wild type strain, the number of *spb1*⁻ bacteria adherent to A549 monolayers was reduced by 60.0% (p<0.01) and the number of intracellular invading bacteria was reduced by 53.6% (p<0.01). Without being bound by any particular theory, it appears that Spb1 may contribute to the pathogenesis of GBS pneumonia and bacterial entry into the bloodstream.

CLAIMS:

1. An isolated nucleic acid molecule comprising nucleotides which code for the amino acid sequence of SEQ ID NO: 2.
2. A recombinant vector comprising the nucleic acid molecule of claim 1.
3. The recombinant vector of claim 2, wherein said recombinant vector is a plasmid.
- 5 4. The recombinant vector of claim 2, wherein said recombinant vector is a prokaryotic or eukaryotic expression vector.
5. The recombinant vector of claim 2, wherein the nucleic acid molecule is operably linked to a heterologous promoter.
6. A host cell comprising the nucleic acid molecule of claim 1.
- 10 7. The host cell of claim 6, wherein the host cell is a eukaryotic host cell.
8. The host cell of claim 6, wherein the host cell is a prokaryotic host cell.
9. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1.
10. A recombinant vector comprising the nucleic acid molecule of claim 9.
- 15 11. The recombinant vector of claim 10, wherein said recombinant vector is a plasmid.
12. The recombinant vector of claim 10, wherein said recombinant vector is a prokaryotic or eukaryotic expression vector.
13. The recombinant vector of claim 10, wherein the nucleic acid molecule is operably linked to a heterologous promoter.
- 20 14. A host cell comprising the nucleic acid molecule of claim 9.
15. The host cell of claim 14, wherein the host cell is a eukaryotic host cell.
16. The host cell of claim 14, wherein the host cell is a prokaryotic host cell.
17. An isolated nucleic acid molecule comprising nucleotides which code for the amino acid sequence of SEQ ID NO: 4
- 25 18. A recombinant vector comprising the nucleic acid molecule of claim 17.
19. The recombinant vector of claim 18, wherein said recombinant vector is a plasmid.
20. The recombinant vector of claim 18, wherein said recombinant vector is a prokaryotic or eukaryotic expression vector.
- 30

21. The recombinant vector of claim 18, wherein the nucleic acid molecule is operably linked to a heterologous promoter.
22. A host cell comprising the nucleic acid molecule of claim 17.
23. The host cell of claim 22, wherein the host cell is a eukaryotic host cell.
24. The host cell of claim 22, wherein the host cell is a prokaryotic host cell.
- 5 25. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 3.
26. A recombinant vector comprising the nucleic acid molecule of claim 25.
27. The recombinant vector of claim 26, wherein said recombinant vector is a plasmid.
- 10 28. The recombinant vector of claim 26, wherein said recombinant vector is a prokaryotic or eukaryotic expression vector.
29. The recombinant vector of claim 26, wherein the nucleic acid molecule is operably linked to a heterologous promoter.
30. A host cell comprising the nucleic acid molecule of claim 25.
- 15 31. The host cell of claim 30, wherein the host cell is a eukaryotic host cell.
32. The host cell of claim 30, wherein the host cell is a prokaryotic host cell.
33. A method of immunizing a mammal against Group B streptococci infection, said method comprising administering to the mammal a vaccine comprising an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 2.
- 20 34. The method of claim 33, wherein the vaccine further comprises an adjuvant.
35. The method of claim 34, wherein the adjuvant comprises alum.
36. The method of claim 33, wherein the vaccine further comprises an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 4.
- 25 37. The method of claim 36, wherein the vaccine further comprises an adjuvant.
38. A method of immunizing a mammal against Group B streptococci infection, said method comprising administering to the mammal a vaccine comprising an immunologically effective amount of a recombinantly produced protein comprising the amino acid sequence of SEQ ID NO: 4.
- 30 39. The method of claim 38, wherein the vaccine further comprises an adjuvant.

40. The method of claim 39, wherein the adjuvant comprises alum.
41. A diagnostic method for determining whether a mammal is infected or colonized by virulent Group B streptococci (GBS), said method comprising the steps of:
- (a) collecting a bodily fluid or culture from the mammal;
 - (b) analyzing the bodily fluid or culture for the presence or absence of one or
5 more gene products specific to type III-3 GBS;
- wherein the presence of one or more gene products specific to type III-3 GBS indicates infection or colonization by virulent GBS.
42. The diagnostic method of claim 41, wherein the one or more gene products specific to type III-3 GBS comprise a protein, said protein comprising the amino acid
10 sequence of SEQ ID NO: 2.
43. The diagnostic method of claim 41, wherein the one or more gene products specific to type III-3 GBS comprise a protein, said protein comprising the amino acid sequence of SEQ ID NO: 4.
44. The diagnostic method of claim 41, wherein the one or more gene products
15 specific to type III-3 GBS comprise a first protein comprising the amino acid sequence of SEQ ID NO: 2 and a second protein comprising the amino acid sequence of SEQ ID NO: 4.
45. The diagnostic method of claim 41, wherein the mammal is a human.
46. The diagnostic method of claim 41, wherein the bodily fluid or culture is a vaginal
20 or rectovaginal culture.
47. The diagnostic method of claim 41, wherein the bodily fluid or culture is a throat culture.
48. The diagnostic method of claim 41, wherein the bodily fluid or culture is blood, serum, amniotic fluid, cerebrospinal fluid, or joint fluid.
- 25 49. The diagnostic method of claim 41, wherein the analysis step comprises using polymerase chain reaction (PCR) to identify the presence or absence of one or more gene products specific to type III-3 GBS.
50. The diagnostic method of claim 49, wherein the analysis step comprises using PCR to determine the presence or absence of the *spb1* gene product.
- 30 51. The diagnostic method of claim 49, wherein the analysis step comprises using PCR to determine the presence or absence of the *spb2* gene product.

52. The diagnostic method of claim 41, wherein the analysis step comprises using antibodies to identify the presence or absence of one or more gene products specific to type III-3 GBS.

53. The diagnostic method of claim 52, wherein the antibodies are monoclonal antibodies.

5 54. The diagnostic method of claim 53, wherein the monoclonal antibodies are specific for the *spb1* gene product.

55. The diagnostic method of claim 53, wherein the monoclonal antibodies are specific for the *spb2* gene product.

10 56. An isolated and purified protein comprising the amino acid sequence of SEQ ID NO: 2.

57. A vaccine for immunizing a mammalian host against virulent Group B streptococci infection, said vaccine comprising the protein of claim 56.

58. The vaccine of claim 57, further comprising an adjuvant.

59. The vaccine of claim 58, wherein the adjuvant comprises alum.

15 60. The vaccine of claim 57, wherein the protein is conjugated to a bacterial polysaccharide or oligosaccharide.

61. An isolated and purified protein comprising the amino acid sequence of SEQ ID NO: 4.

20 62. A vaccine for immunizing a mammalian host against virulent Group B streptococci infection, said vaccine comprising the protein of claim 61.

63. The vaccine of claim 62, further comprising an adjuvant.

64. The vaccine of claim 63, wherein the adjuvant comprises alum.

65. The vaccine of claim 62, further comprising an isolated and purified protein comprising the amino acid sequence of SEQ ID NO: 2.

25 66. The vaccine of claim 65, further comprising an adjuvant.

67. The vaccine of claim 62, wherein the protein is conjugated to a bacterial polysaccharide or oligosaccharide.

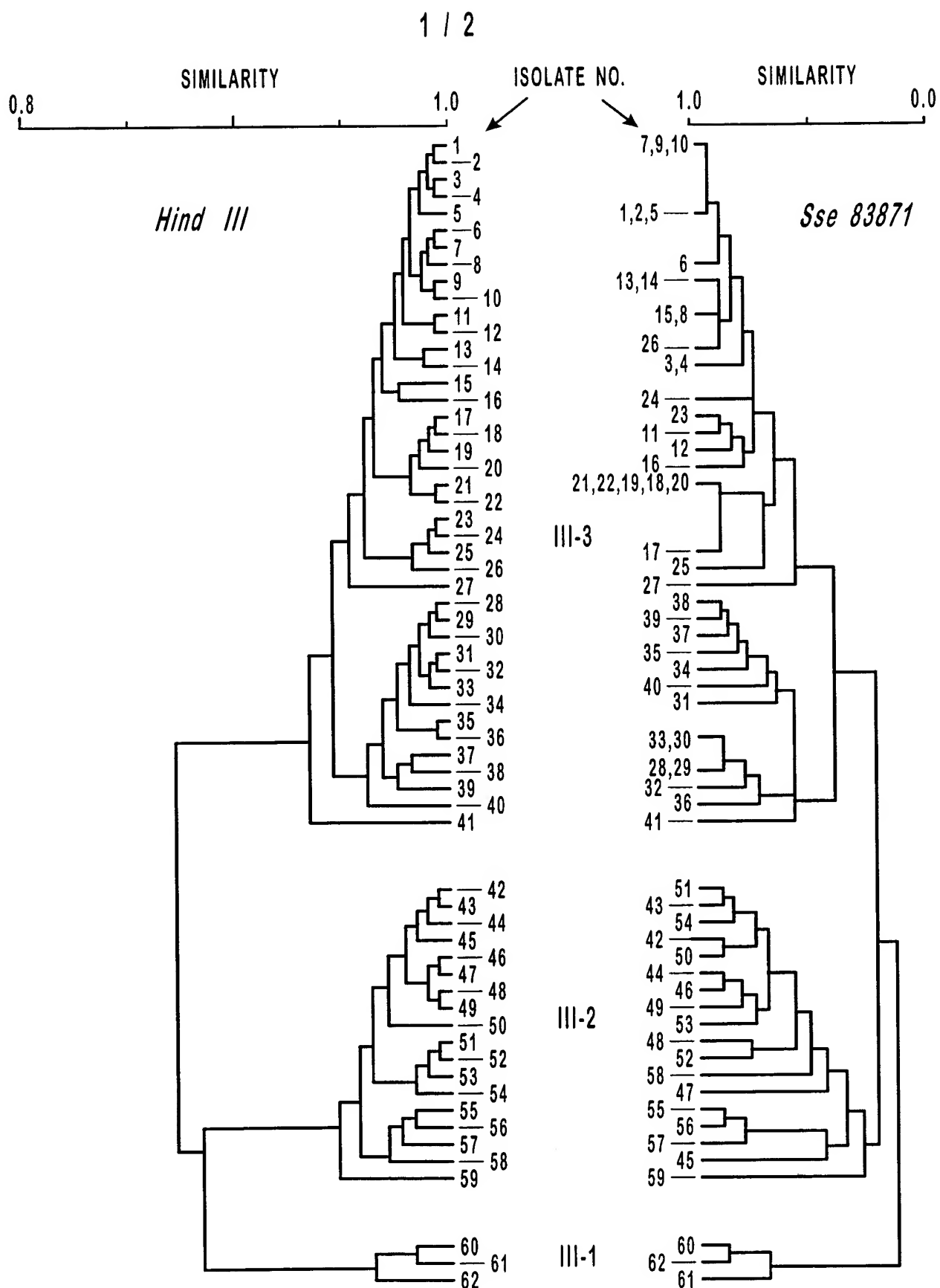


FIG. 1

2 / 2

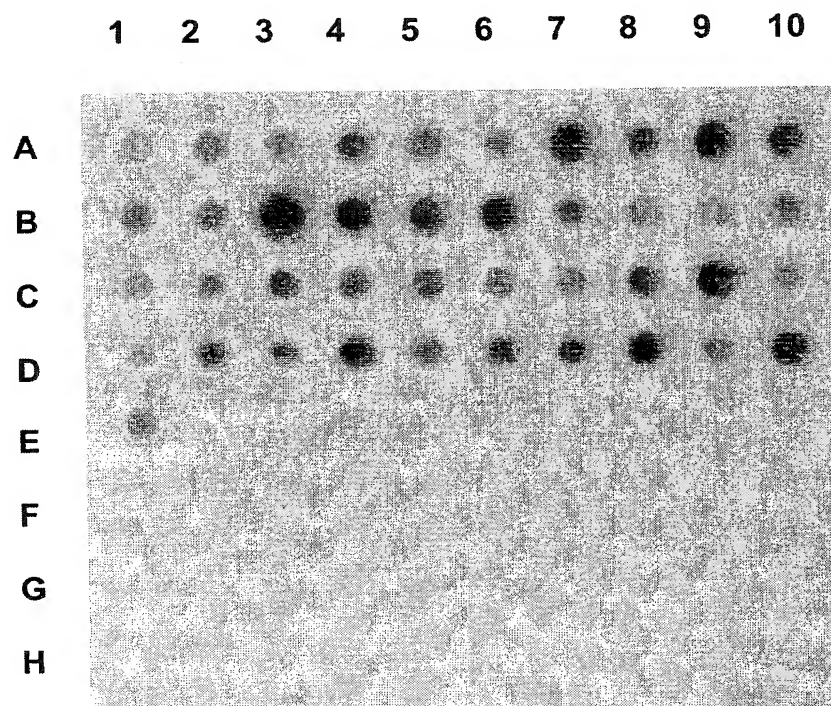


FIG. 2

SEQUENCE LISTING

<110> Adderson, Elisabeth E.
 Bohnsack, John F.
 University of Utah Research Foundation

<120> Isolated Genes from Virulent Group B Streptococcus
 agalactiae

<130> 1321.2.29

<140>
 <141>

<150> 60/140,084
 <151> 1999-06-21

<160> 10

<170> PatentIn Ver. 2.1

<210> 1
 <211> 1509
 <212> DNA
 <213> Streptococcus agalactiae

<220>
 <221> CDS
 <222> (1)..(1509)

<400> 1
 atg aaa aag aaa atg att caa tcg ctg tta gtg gcg agt tta gca ttt 48
 Met Lys Lys Lys Met Ile Gln Ser Leu Leu Val Ala Ser Leu Ala Phe
 1 5 10 15
 ggt atg gct gta tca cca gtt acg ccg ata gct ttt gcc gct gag aca 96
 Gly Met Ala Val Ser Pro Val Thr Pro Ile Ala Phe Ala Ala Glu Thr
 20 25 30
 ggg aca att aca gtt caa gat act caa aaa ggc gca acc tat aaa gca 144
 Gly Thr Ile Thr Val Gln Asp Thr Gln Lys Gly Ala Thr Tyr Lys Ala
 35 40 45
 tat aaa gtt ttt gat gca gaa ata gat aat gca aat gta tct gat tcg 192
 Tyr Lys Val Phe Asp Ala Glu Ile Asp Asn Ala Asn Val Ser Asp Ser
 50 55 60
 aat aaa gat gga gct tct tat tta att cct caa ggt aaa gaa gct gag 240

2

Gly	Ser	Glu	Lys	Ala	Thr	Gly	Lys	Tyr	Asn	Leu	Leu	Glu	Glu	Asn	Asn		
			260					265					270				
aat	ttc	acg	att	act	att	ccg	tgg	gca	gct	acc	aat	act	cca	acc	gga	864	
Asn	Phe	Thr	Ile	Thr	Ile	Pro	Trp	Ala	Ala	Thr	Asn	Thr	Pro	Thr	Gly		
		275					280					285					
aat	act	caa	aat	gga	gct	aat	gat	gac	ttt	ttt	tat	aag	gga	ata	aat	912	
Asn	Thr	Gln	Asn	Gly	Ala	Asn	Asp	Asp	Phe	Phe	Tyr	Lys	Gly	Ile	Asn		
	290						295				300						
aca	atc	aca	gtc	act	tat	aca	gga	gta	tta	aag	agt	gga	gct	aaa	cca	960	
Thr	Ile	Thr	Val	Thr	Tyr	Thr	Gly	Val	Leu	Lys	Ser	Gly	Ala	Lys	Pro		
305					310					315					320		
ggt	tca	gct	gat	tta	cca	gaa	aat	aca	aac	att	gcg	acc	atc	aac	ccc	1008	
Gly	Ser	Ala	Asp	Leu	Pro	Glu	Asn	Thr	Asn	Ile	Ala	Thr	Ile	Asn	Pro		
			325						330					335			
aat	act	agc	aat	gat	gac	cca	ggt	caa	aaa	gta	aca	gtg	agg	gat	ggt	1056	
Asn	Thr	Ser	Asn	Asp	Asp	Pro	Gly	Gln	Lys	Val	Thr	Val	Arg	Asp	Gly		
		340						345					350				
caa	att	act	ata	aaa	aaa	att	gat	ggt	tcc	aca	aaa	gct	tca	tta	caa	1104	
Gln	Ile	Thr	Ile	Lys	Lys	Ile	Asp	Gly	Ser	Thr	Lys	Ala	Ser	Leu	Gln		
		355					360					365					
ggt	gct	ata	ttt	gtt	tta	aag	aat	gct	acg	ggt	caa	ttt	cta	aac	ttt	1152	
Gly	Ala	Ile	Phe	Val	Leu	Lys	Asn	Ala	Thr	Gly	Gln	Phe	Leu	Asn	Phe		
	370					375					380						
aac	gat	aca	aat	aac	gtt	gaa	tgg	ggc	aca	gaa	gct	aat	gca	aca	gaa	1200	
Asn	Asp	Thr	Asn	Asn	Val	Glu	Trp	Gly	Thr	Glu	Ala	Asn	Ala	Thr	Glu		
385					390					395					400		
tat	aca	aca	gga	gca	gat	ggt	ata	att	acc	att	aca	ggc	ttg	aaa	gaa	1248	
Tyr	Thr	Thr	Gly	Ala	Asp	Gly	Ile	Ile	Thr	Ile	Thr	Gly	Leu	Lys	Glu		
			405					410					415				
ggt	aca	tac	tat	cta	gtt	gag	aaa	aag	gct	ccc	tta	ggt	tac	aat	ttg	1296	
Gly	Thr	Tyr	Tyr	Leu	Val	Glu	Lys	Lys	Ala	Pro	Leu	Gly	Tyr	Asn	Leu		
		420						425					430				
tta	gat	aac	tct	cag	aag	gtt	att	tta	gga	gat	gga	gcc	act	gat	acg	1344	
Leu	Asp	Asn	Ser	Gln	Lys	Val	Ile	Leu	Gly	Asp	Gly	Ala	Thr	Asp	Thr		
		435					440					445					
act	aat	tca	gat	aac	ctt	tta	gtt	aac	cca	act	gtt	gaa	aat	aac	aaa	1392	

Thr Asn Ser Asp Asn Leu Leu Val Asn Pro Thr Val Glu Asn Asn Lys
 450 455 460

ggt act gag ttg cct tca aca ggt ggt att ggt aca aca att ttc tac 1440
 Gly Thr Glu Leu Pro Ser Thr Gly Gly Ile Gly Thr Thr Ile Phe Tyr
 465 470 475 480

att ata ggt gca att tta gta ata gga gca ggt atc gtg ctt gtt gct 1488
 Ile Ile Gly Ala Ile Leu Val Ile Gly Ala Gly Ile Val Leu Val Ala
 485 490 495

cgt cgt cgt tta cgt tct taa 1509
 Arg Arg Arg Leu Arg Ser
 500

<210> 2

<211> 502

<212> PRT

<213> Streptococcus agalactiae

<400> 2

Met Lys Lys Lys Met Ile Gln Ser Leu Leu Val Ala Ser Leu Ala Phe
 1 5 10 15
 Gly Met Ala Val Ser Pro Val Thr Pro Ile Ala Phe Ala Ala Glu Thr
 20 25 30
 Gly Thr Ile Thr Val Gln Asp Thr Gln Lys Gly Ala Thr Tyr Lys Ala
 35 40 45
 Tyr Lys Val Phe Asp Ala Glu Ile Asp Asn Ala Asn Val Ser Asp Ser
 50 55 60
 Asn Lys Asp Gly Ala Ser Tyr Leu Ile Pro Gln Gly Lys Glu Ala Glu
 65 70 75 80
 Tyr Lys Ala Ser Thr Asp Phe Asn Ser Leu Phe Thr Thr Thr Thr Asn
 85 90 95
 Gly Gly Arg Thr Tyr Val Thr Lys Lys Asp Thr Ala Ser Ala Asn Glu
 100 105 110
 Ile Ala Thr Trp Ala Lys Ser Ile Ser Ala Asn Thr Thr Pro Val Ser
 115 120 125
 Thr Val Thr Glu Ser Asn Asn Asp Gly Thr Glu Val Ile Asn Val Ser
 130 135 140
 Gln Tyr Gly Tyr Tyr Tyr Val Ser Ser Thr Val Asn Asn Gly Ala Val
 145 150 155 160
 Ile Met Val Thr Ser Val Thr Pro Asn Ala Thr Ile His Glu Lys Asn
 165 170 175
 Thr Asp Ala Thr Trp Gly Asp Gly Gly Gly Lys Thr Val Asp Gln Lys
 180 185 190
 Thr Tyr Ser Val Gly Asp Thr Val Lys Tyr Thr Ile Thr Tyr Lys Asn
 195 200 205

Ala Val Asn Tyr His Gly Thr Glu Lys Val Tyr Gln Tyr Val Ile Lys
 210 215 220
 Asp Thr Met Pro Ser Ala Ser Val Val Asp Leu Asn Glu Gly Ser Tyr
 225 230 235 240
 Glu Val Thr Ile Thr Asp Gly Ser Gly Asn Ile Thr Thr Leu Thr Gln
 245 250 255
 Gly Ser Glu Lys Ala Thr Gly Lys Tyr Asn Leu Leu Glu Glu Asn Asn
 260 265 270
 Asn Phe Thr Ile Thr Ile Pro Trp Ala Ala Thr Asn Thr Pro Thr Gly
 275 280 285
 Asn Thr Gln Asn Gly Ala Asn Asp Asp Phe Phe Tyr Lys Gly Ile Asn
 290 295 300
 Thr Ile Thr Val Thr Tyr Thr Gly Val Leu Lys Ser Gly Ala Lys Pro
 305 310 315 320
 Gly Ser Ala Asp Leu Pro Glu Asn Thr Asn Ile Ala Thr Ile Asn Pro
 325 330 335
 Asn Thr Ser Asn Asp Asp Pro Gly Gln Lys Val Thr Val Arg Asp Gly
 340 345 350
 Gln Ile Thr Ile Lys Lys Ile Asp Gly Ser Thr Lys Ala Ser Leu Gln
 355 360 365
 Gly Ala Ile Phe Val Leu Lys Asn Ala Thr Gly Gln Phe Leu Asn Phe
 370 375 380
 Asn Asp Thr Asn Asn Val Glu Trp Gly Thr Glu Ala Asn Ala Thr Glu
 385 390 395 400
 Tyr Thr Thr Gly Ala Asp Gly Ile Ile Thr Ile Thr Gly Leu Lys Glu
 405 410 415
 Gly Thr Tyr Tyr Leu Val Glu Lys Lys Ala Pro Leu Gly Tyr Asn Leu
 420 425 430
 Leu Asp Asn Ser Gln Lys Val Ile Leu Gly Asp Gly Ala Thr Asp Thr
 435 440 445
 Thr Asn Ser Asp Asn Leu Leu Val Asn Pro Thr Val Glu Asn Asn Lys
 450 455 460
 Gly Thr Glu Leu Pro Ser Thr Gly Gly Ile Gly Thr Thr Ile Phe Tyr
 465 470 475 480
 Ile Ile Gly Ala Ile Leu Val Ile Gly Ala Gly Ile Val Leu Val Ala
 485 490 495
 Arg Arg Arg Leu Arg Ser
 500

<210> 3

<211> 1692

<212> DNA

<213> Streptococcus agalactiae

<220>

<221> CDS

<222> (1) .. (1692)

<400> 3

```

att tgc att atg gtg atc gta ttc cgg att ata cag ata tta caa ggg      48
Ile Cys Ile Met Val Ile Val Phe Arg Ile Ile Gln Ile Leu Gln Gly
  1               5               10               15

att ata tcc aag atc ctt cag gta cat att att ata agt atg att cac      96
Ile Ile Ser Lys Ile Leu Gln Val His Ile Ile Ile Ser Met Ile His
      20               25               30

gag ata aag atc ccg act caa cta aag atg cct att ata cga cag ata     144
Glu Ile Lys Ile Pro Thr Gln Leu Lys Met Pro Ile Ile Arg Gln Ile
      35               40               45

cta gtc tca tca aat gtt gat aca aca act aag tac aag tac gta aaa     192
Leu Val Ser Ser Asn Val Asp Thr Thr Thr Lys Tyr Lys Tyr Val Lys
      50               55               60

gac gct tac aaa tta gtc ggt tgg tat tat gtt aat cca tat ggt agt     240
Asp Ala Tyr Lys Leu Val Gly Trp Tyr Tyr Val Asn Pro Tyr Gly Ser
      65               70               75               80

att aga cct tat aac ttt tca ggt gct gta act caa gat atc aat tta     288
Ile Arg Pro Tyr Asn Phe Ser Gly Ala Val Thr Gln Asp Ile Asn Leu
      85               90               95

aga gct att tgg cga aag gct gga gat tat cat att ata tac agc aat     336
Arg Ala Ile Trp Arg Lys Ala Gly Asp Tyr His Ile Ile Tyr Ser Asn
      100              105              110

gat gct gtt ggt aca gat gga aag cca gca ttg gat gct tct ggt cag     384
Asp Ala Val Gly Thr Asp Gly Lys Pro Ala Leu Asp Ala Ser Gly Gln
      115              120              125

caa tta caa aca agt aat gag cct act gac cct gat tcc tat gac gat     432
Gln Leu Gln Thr Ser Asn Glu Pro Thr Asp Pro Asp Ser Tyr Asp Asp
      130              135              140

ggc tcc cat tca gcc tta ctg aga cgt ccg aca atg cca gat ggc tat     480
Gly Ser His Ser Ala Leu Leu Arg Arg Pro Thr Met Pro Asp Gly Tyr
      145              150              155              160

cgt ttc cgt ggc tgg tgg tac aat ggt aaa att tat aac cca tat gat     528
Arg Phe Arg Gly Trp Trp Tyr Asn Gly Lys Ile Tyr Asn Pro Tyr Asp
      165              170              175

tcc att gat att gac gcc cat tta gca gat gct aat aaa aat atc acc     576

```

Ser Ile Asp Ile Asp Ala His Leu Ala Asp Ala Asn Lys Asn Ile Thr	
180 185 190	
ata aaa cct gtc att att cca gta gga gat atc aaa tta gaa gat acc	624
Ile Lys Pro Val Ile Ile Pro Val Gly Asp Ile Lys Leu Glu Asp Thr	
195 200 205	
tcc atc aaa tac aat ggt aac ggt ggt act aga gta gaa aat ggt aat	672
Ser Ile Lys Tyr Asn Gly Asn Gly Gly Thr Arg Val Glu Asn Gly Asn	
210 215 220	
gtg gta aca caa gtg gag aca ccg cgt atg gag ttg aat agc aca act	720
Val Val Thr Gln Val Glu Thr Pro Arg Met Glu Leu Asn Ser Thr Thr	
225 230 235 240	
aca att cct gaa aac caa tac ttt aca agg aca ggt tac aac ctt att	768
Thr Ile Pro Glu Asn Gln Tyr Phe Thr Arg Thr Gly Tyr Asn Leu Ile	
245 250 255	
ggt tgg cat cat gat aag gat tta gct gat aca gga cgt gtg gaa ttt	816
Gly Trp His His Asp Lys Asp Leu Ala Asp Thr Gly Arg Val Glu Phe	
260 265 270	
aca gca ggt caa tca ata ggt att gat aac aac ctt gat gca aca aat	864
Thr Ala Gly Gln Ser Ile Gly Ile Asp Asn Asn Leu Asp Ala Thr Asn	
275 280 285	
acc tta tat gct gtt tgg caa cct aaa gaa tac acc gtc gga gta agt	912
Thr Leu Tyr Ala Val Trp Gln Pro Lys Glu Tyr Thr Val Gly Val Ser	
290 295 300	
aaa act gtc gtt gga cta gat gaa gat aag acg aaa gac ttc ttg ttt	960
Lys Thr Val Val Gly Leu Asp Glu Asp Lys Thr Lys Asp Phe Leu Phe	
305 310 315 320	
aat cca agt gaa acg ttg caa caa gag aat ttt ccg ctg aga gat ggt	1008
Asn Pro Ser Glu Thr Leu Gln Gln Glu Asn Phe Pro Leu Arg Asp Gly	
325 330 335	
cag act aag gaa ttt aaa gta cct tat gga act tct ata tca ata gat	1056
Gln Thr Lys Glu Phe Lys Val Pro Tyr Gly Thr Ser Ile Ser Ile Asp	
340 345 350	
gaa caa gcc tac gat gaa ttt aaa gta tct gag tca att aca gaa aaa	1104
Glu Gln Ala Tyr Asp Glu Phe Lys Val Ser Glu Ser Ile Thr Glu Lys	
355 360 365	
aat cta gca act ggt gaa gct gat aaa act tat gat gct acc ggc tta	1152

Asn Leu Ala Thr Gly Glu Ala Asp Lys Thr Tyr Asp Ala Thr Gly Leu	
370 375 380	
caa tcc ctg aca gtt tca gga gac gta gat att agc ttt acc aat aca	1200
Gln Ser Leu Thr Val Ser Gly Asp Val Asp Ile Ser Phe Thr Asn Thr	
385 390 395 400	
cgt atc aag caa aaa gta cga cta cag aaa gtt aat gtc gaa aat gat	1248
Arg Ile Lys Gln Lys Val Arg Leu Gln Lys Val Asn Val Glu Asn Asp	
405 410 415	
aat aat ttt tta gca ggt gca gtt ttt gat att tat gaa tca gat gct	1296
Asn Asn Phe Leu Ala Gly Ala Val Phe Asp Ile Tyr Glu Ser Asp Ala	
420 425 430	
aat ggg aat aaa gct tca cat cct atg tat tca ggg ctg gtg aca aac	1344
Asn Gly Asn Lys Ala Ser His Pro Met Tyr Ser Gly Leu Val Thr Asn	
435 440 445	
gat aaa ggc ttg tta tta gtg gat gct aat aac tac ctc agt ttg cca	1392
Asp Lys Gly Leu Leu Leu Val Asp Ala Asn Asn Tyr Leu Ser Leu Pro	
450 455 460	
gta gga aaa tac tac cta aca gag aca aag gcc cct cca ggg tac cta	1440
Val Gly Lys Tyr Tyr Leu Thr Glu Thr Lys Ala Pro Pro Gly Tyr Leu	
465 470 475 480	
cta cct aaa aat gat gat ata tca gta tta gtg att tct acg gga gtt	1488
Leu Pro Lys Asn Asp Asp Ile Ser Val Leu Val Ile Ser Thr Gly Val	
485 490 495	
acc ttt gaa caa aat ggt aat aat gcg aca cca ata aaa gag aat tta	1536
Thr Phe Glu Gln Asn Gly Asn Asn Ala Thr Pro Ile Lys Glu Asn Leu	
500 505 510	
gtg gat gga agt aca gta tat act ttt aaa att act aac agt aaa gga	1584
Val Asp Gly Ser Thr Val Tyr Thr Phe Lys Ile Thr Asn Ser Lys Gly	
515 520 525	
aca gaa ttg cct agt act gga ggt att gga aca cac att tat atc cta	1632
Thr Glu Leu Pro Ser Thr Gly Gly Ile Gly Thr His Ile Tyr Ile Leu	
530 535 540	
gtt ggt tta gct tta gct cta cca tca gga tta ata tta tac tat cga	1680
Val Gly Leu Ala Leu Ala Leu Pro Ser Gly Leu Ile Leu Tyr Tyr Arg	
545 550 555 560	
aaa aaa ata tga	1692

Lys Lys Ile

<210> 4

<211> 563

<212> PRT

<213> Streptococcus agalactiae

<400> 4

```

Ile Cys Ile Met Val Ile Val Phe Arg Ile Ile Gln Ile Leu Gln Gly
 1           5           10           15
Ile Ile Ser Lys Ile Leu Gln Val His Ile Ile Ile Ser Met Ile His
          20           25           30
Glu Ile Lys Ile Pro Thr Gln Leu Lys Met Pro Ile Ile Arg Gln Ile
          35           40           45
Leu Val Ser Ser Asn Val Asp Thr Thr Thr Lys Tyr Lys Tyr Val Lys
          50           55           60
Asp Ala Tyr Lys Leu Val Gly Trp Tyr Tyr Val Asn Pro Tyr Gly Ser
 65           70           75           80
Ile Arg Pro Tyr Asn Phe Ser Gly Ala Val Thr Gln Asp Ile Asn Leu
          85           90           95
Arg Ala Ile Trp Arg Lys Ala Gly Asp Tyr His Ile Ile Tyr Ser Asn
          100          105          110
Asp Ala Val Gly Thr Asp Gly Lys Pro Ala Leu Asp Ala Ser Gly Gln
          115          120          125
Gln Leu Gln Thr Ser Asn Glu Pro Thr Asp Pro Asp Ser Tyr Asp Asp
          130          135          140
Gly Ser His Ser Ala Leu Leu Arg Arg Pro Thr Met Pro Asp Gly Tyr
145           150           155           160
Arg Phe Arg Gly Trp Trp Tyr Asn Gly Lys Ile Tyr Asn Pro Tyr Asp
          165          170          175
Ser Ile Asp Ile Asp Ala His Leu Ala Asp Ala Asn Lys Asn Ile Thr
          180          185          190
Ile Lys Pro Val Ile Ile Pro Val Gly Asp Ile Lys Leu Glu Asp Thr
          195          200          205
Ser Ile Lys Tyr Asn Gly Asn Gly Gly Thr Arg Val Glu Asn Gly Asn
          210          215          220
Val Val Thr Gln Val Glu Thr Pro Arg Met Glu Leu Asn Ser Thr Thr
225           230           235           240
Thr Ile Pro Glu Asn Gln Tyr Phe Thr Arg Thr Gly Tyr Asn Leu Ile
          245          250          255
Gly Trp His His Asp Lys Asp Leu Ala Asp Thr Gly Arg Val Glu Phe
          260          265          270
Thr Ala Gly Gln Ser Ile Gly Ile Asp Asn Asn Leu Asp Ala Thr Asn
          275          280          285
Thr Leu Tyr Ala Val Trp Gln Pro Lys Glu Tyr Thr Val Gly Val Ser
          290          295          300
Lys Thr Val Val Gly Leu Asp Glu Asp Lys Thr Lys Asp Phe Leu Phe

```

```

305          310          315          320
Asn Pro Ser Glu Thr Leu Gln Gln Glu Asn Phe Pro Leu Arg Asp Gly
          325          330          335
Gln Thr Lys Glu Phe Lys Val Pro Tyr Gly Thr Ser Ile Ser Ile Asp
          340          345          350
Glu Gln Ala Tyr Asp Glu Phe Lys Val Ser Glu Ser Ile Thr Glu Lys
          355          360          365
Asn Leu Ala Thr Gly Glu Ala Asp Lys Thr Tyr Asp Ala Thr Gly Leu
          370          375          380
Gln Ser Leu Thr Val Ser Gly Asp Val Asp Ile Ser Phe Thr Asn Thr
385          390          395          400
Arg Ile Lys Gln Lys Val Arg Leu Gln Lys Val Asn Val Glu Asn Asp
          405          410          415
Asn Asn Phe Leu Ala Gly Ala Val Phe Asp Ile Tyr Glu Ser Asp Ala
          420          425          430
Asn Gly Asn Lys Ala Ser His Pro Met Tyr Ser Gly Leu Val Thr Asn
          435          440          445
Asp Lys Gly Leu Leu Leu Val Asp Ala Asn Asn Tyr Leu Ser Leu Pro
          450          455          460
Val Gly Lys Tyr Tyr Leu Thr Glu Thr Lys Ala Pro Pro Gly Tyr Leu
465          470          475          480
Leu Pro Lys Asn Asp Asp Ile Ser Val Leu Val Ile Ser Thr Gly Val
          485          490          495
Thr Phe Glu Gln Asn Gly Asn Asn Ala Thr Pro Ile Lys Glu Asn Leu
          500          505          510
Val Asp Gly Ser Thr Val Tyr Thr Phe Lys Ile Thr Asn Ser Lys Gly
          515          520          525
Thr Glu Leu Pro Ser Thr Gly Gly Ile Gly Thr His Ile Tyr Ile Leu
          530          535          540
Val Gly Leu Ala Leu Ala Leu Pro Ser Gly Leu Ile Leu Tyr Tyr Arg
545          550          555          560
Lys Lys Ile

```

<210> 5

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic
oligonucleotide

<400> 5

ctaggtggat ccttcggcaa t

21

<210> 6
<211> 10
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic
oligonucleotide

<400> 6
cgattgccga

10

<210> 7
<211> 25
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic
oligonucleotide

<400> 7
aggcaactgt gctaaccgag ggaat

25

<210> 8
<211> 11
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic
oligonucleotide

<400> 8
cgattccctc g

11

<210> 9
<211> 5
<212> PRT
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Consensus
sequence

<400> 9

Leu Pro Xaa Thr Gly

1 5

<210> 10

<211> 6

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Consensus
sequence

<400> 10

Leu Pro Ser Thr Gly Gly

1 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/17082

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : Please See Extra Sheet.

US CL : Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.7; 530/350, 825; 424/244.1, 184.1, 236.1, 234.1, 193.1; 435/320.1, 69.3, 71.1, 7.1, 7.34

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

DIALOG, MEDLINE, BIOSIS, EMBASE, WEST

spb1, spb2, GBS, SEQ ID NOs: 1, 2, 3 and 4, streptococc?, inventors' names

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P ----- Y,P	ADDERSON et al. Genetic subtraction identifies a novel respiratory epithelial adhesin/invasin of type III group B streptococcus (GBS). Pediatr. Res. 30 March 2000, Vol. 47, No. 4, page 337A, abstract 1992, see entire abstract.	1-16, 56 ----- 17-32, 33-55, 57-67

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
O document referring to an oral disclosure, use, exhibition or other means	
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

31 OCTOBER 2000

Date of mailing of the international search report

14 NOV 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

S. DEVI, Ph.D.

Telephone No. (703) 308-1235

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/17082

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☒ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/17082

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

C07H 21/04; C07K 1/00, A61K 39/09, 39/38, 39/39, 38/385; C12N 15/00, 15/09; C12P 21/04, G01N 33/53, 33/569

A. CLASSIFICATION OF SUBJECT MATTER:

US CL :

536/23.7; 530/350, 825; 424/244.1, 184.1, 236.1, 234.1, 193.1; 435/320.1, 69.3, 71.1, 7.1, 7.34

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

Group I, claim(s) 1-16, 33-35 and 56-60, drawn to an isolated nucleic acid molecule encoding SEQ ID NO: 2, a vector and a host cell comprising the same and a method of immunizing a mammal by administering a protein comprising SEQ ID NO: 2.

Group II, claim(s) 17-32, 38-40, 61-64 and 67, drawn to an isolated nucleic acid molecule encoding SEQ ID NO: 4, a vector and a host cell comprising the same and a method of immunizing a mammal by administering a protein comprising SEQ ID NO: 4.

Group III, claim(s) 42, 50 and 54, drawn to a diagnostic method comprising analyzing for the presence or absence of SEQ ID NO: 2.

Group IV, claims 43, 51 and 55, drawn to a diagnostic method comprising analyzing for the presence or absence of SEQ ID NO: 4.

Claims 36, 37, 65 and 66 are considered linking claims and would be joined with Group I or II, if elected.

Claims 41, 44-49, 52 and 53 are considered linking claims and would be joined with Group II or III, if elected.

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Group I is directed to a nucleic acid molecule of SEQ ID NO:1, a protein with amino acid sequence of SEQ ID NO: 2, a vector and a host cell comprising the nucleic acid molecule and the first method of use of the protein in immunization. This is a permitted category under PCT Rule 13.2. Group II is directed to a second product, a nucleic acid molecule of SEQ ID NO:3, protein with amino acid sequence of SEQ ID NO: 4, a vector and a host cell comprising the nucleic acid molecule and a method of immunizing a mammal by administering the protein. Groups III and IV are directed respectively to a second method of use of the gene products, i.e., diagnostic methods for analyzing spb1 and spb2 gene products respectively. The SEQ ID NO.s of inventions I and II are the special technical features, which distinct from one another in their structure and specificity. The inventions therefore lack unity because the special technical feature is not a unifying feature.